GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 03:11:06 ; Search time 918.465 Seconds
(without alignments)
9617.545 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcggcacgagcggcac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1400	96.2	1755	18	US-10-425-115-18040	Sequence 18040, A
	2	1394.2	95.8	1444	17	US-10-425-114-15440	Sequence 15440, A
	3	1394.2	95.8	1450	17	US-10-425-114-25244	Sequence 25244, A
	4	1196.8	82.3	1706	18	US-10-425-115-18039	Sequence 18039, A
	5	1179	81.0	1649	17	US-10-425-114-24871	Sequence 24871, A
	6	1113.4	76.5	1537	18	US-10-767-701-13951	Sequence 13951, A
	7	1047	72.0	1361	17	US-10-425-114-25845	Sequence 25845, A
	8	1047	72.0	1490	17	US-10-425-114-22598	Sequence 22598, A
	9	1046	71.9	1519	18	US-10-425-115-18041	Sequence 18041, A
c	10	753	51.8	2471	18	US-10-739-930-4858	Sequence 4858, Ap
	11	704.8	48.4	1569	18	US-10-437-963-37229	Sequence 37229, A
	12	571	39.2	571	16	US-10-074-473-2	Sequence 2, Appli
	13	479	32.9	1691	17	US-10-424-599-65699	Sequence 65699, A
	14	476.2	32.7	1598	17	US-10-424-599-81927	Sequence 81927, A
	15	416.6	28.6	483	17	US-10-424-599-118531	Sequence 118531,
	16	388	26.7	388	16	US-10-074-473-3	Sequence 3, Appli
	17	332.8	22.9	1674	17	US-10-424-599-98272	Sequence 98272, A
	18	286	19.7	287	9	US-09-294-093B-571	Sequence 571, App
	19	245.2	16.9	1029	17	US-10-424-599-81933	Sequence 81933, A
c	20	227	15.6	316	18	US-10-425-115-87023	Sequence 87023, A
c	21	223.4	15.4	384	18	US-10-425-115-41982	Sequence 41982, A
	22	221.8	15.2	585	18	US-10-021-323-2492	Sequence 2492, Ap
	23	217.8	15.0	534	18	US-10-021-323-14417	Sequence 14417, A
	24	216.4	14.9	272	9	US-09-294-093B-4704	Sequence 4704, Ap
	25	201.6	13.9	291	9	US-09-294-093B-3135	Sequence 3135, Ap
	26	201.6	13.9	851	17	US-10-424-599-75856	Sequence 75856, A
	27	194.4	13.4	528	18	US-10-021-323-14351	Sequence 14351, A
c	28	186.2	12.8	329	18	US-10-425-115-137486	Sequence 137486,
c	29	183.8	12.6	573	18	US-10-021-323-2546	Sequence 2546, Ap
	30	182.8	12.6	275	18	US-10-425-115-60448	Sequence 60448, A
	31	180.6	12.4	656	17	US-10-424-599-98270	Sequence 98270, A
c	32	175.6	12.1	523	18	US-10-021-323-15961	Sequence 15961, A
	33	163	11.2	1555	18	US-10-335-053-99	Sequence 99, Appl
	34	163	11.2	1776	18	US-10-357-930-22330	Sequence 22330, A
	35	163	11.2	1776	18	US-10-357-930-28174	Sequence 28174, A
	36	156.4	10.7	1264	17	US-10-264-049-918	Sequence 918, App
	37	156.4	10.7	1366	16	US-10-287-218-33	Sequence 33, Appl
	38	156.4	10.7	1366	18	US-10-474-291-33	Sequence 33, Appl
	39	156.4	10.7	2053	10	US-09-983-802-45	Sequence 45, Appl
	40	156.4	10.7	2053	10	US-09-984-490-45	Sequence 45, Appl
	41	156.4	10.7	2053	11	US-09-973-278-82	Sequence 82, Appl
c	42	142.2	9.8	409	9	US-09-770-423-621	Sequence 621, App
	43	129.2	8.9	481	10	US-09-770-961-929	Sequence 929, App
c	44	108.2	7.4	432	17	US-10-424-599-49636	Sequence 49636, A
c	45	106.8	7.3	479	17	US-10-424-599-25942	Sequence 25942, A

ALIGNMENTS

RESULT 1
US-10-425-115-18040
; Sequence 18040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18040
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116455C.1
US-10-425-115-18040

Query Match 96.2%; Score 1400; DB 18; Length 1755;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1412; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         194 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 253

Qy          84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         254 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 313
```

```
Qy      144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
Db      314 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 373

Qy      204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
Db      374 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 433

Qy      264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
Db      434 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 493

Qy      324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
Db      494 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 553

Qy      384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
Db      554 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 613

Qy      444 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
Db      614 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 673

Qy      504 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 563
Db      674 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 733

Qy      564 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 623
Db      734 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 793

Qy      624 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
Db      794 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 853

Qy      684 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 743
Db      854 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 913

Qy      744 GAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCA 803
Db      914 GAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCA 973

Qy      804 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 863
Db      974 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 1033

Qy      864 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 923
Db     1034 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 1093

Qy      924 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 983
Db     1094 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 1153

Qy      984 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1043
Db     1154 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1213

Qy     1044 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1103
Db     1214 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1273

Qy     1104 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1163
Db     1274 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1333

Qy     1164 CTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCT 1223
Db     1334 CTAAAGTTTGGACTGCCCATCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCC 1393

Qy     1224 GCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1283
Db     1394 GTCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1453

Qy     1284 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1343
Db     1454 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1513

Qy     1344 TATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1403
Db     1514 TATTATGTGGTATCGACCGGGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1573

Qy     1404 GGAATGCCTATTCTTTTTGGTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1455
Db     1574 GGAATGCCTATTCTTTTTGGTCTAAGATGTTTTTCTTGTCAAAAAAAAAAAA 1625
```

RESULT 2
US-10-425-114-15440
; Sequence 15440, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15440
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-061-D10_FLI
US-10-425-114-15440

Query Match 95.8%; Score 1394.2; DB 17; Length 1444;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy       24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
Db       25 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAACCGGAAGGGGCGCCAT 84

Qy       84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
Db       85 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 144

Qy      144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
Db      145 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 204

Qy      204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
Db      205 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 264

Qy      264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
Db      265 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 324

Qy      324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
Db      325 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 384

Qy      384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
Db      385 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 444
```

```
Qy        444 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 504

Qy        504 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 564

Qy        564 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 624

Qy        624 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 684

Qy        684 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 744

Qy        744 GAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCA 803
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        745 GAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCA 804

Qy        804 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 864

Qy        864 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 924

Qy        924 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 984

Qy        984 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1044

Qy       1044 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1104

Qy       1104 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1164

Qy       1164 CTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCT 1223
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1165 CTAAAGTTTGGACTGCCCATCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCC 1224

Qy       1224 GCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1283
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1225 GTCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1284

Qy       1284 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1344

Qy       1344 TATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1403
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db       1345 TATTATGTGGTATCGACCGGGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1404

Qy       1404 GGAATGCCTATTCTTTTTGGTCTAAAA 1430
              |||||||||||||||||||||||| |
Db       1405 GGAATGCCTATTCTTTTTGGTCTAAGA 1431
```

RESULT 3
US-10-425-114-25244
; Sequence 25244, Application US/10425114

; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25244
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3732-024-C2_FLI
US-10-425-114-25244

Query Match 95.8%; Score 1394.2; DB 17; Length 1450;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy         24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAACCGGAAGGGGCGCCAT 90

Qy         84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 150

Qy        144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 210

Qy        204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 270

Qy        264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 330

Qy        324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 390

Qy        384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 450

Qy        444 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 510

Qy        504 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 570

Qy        564 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 630

Qy        624 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 690

Qy        684 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 750
```

```
Qy        744 GAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCA 803
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        751 GAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCA 810

Qy        804 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 870

Qy        864 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 930

Qy        924 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 990

Qy        984 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1050

Qy       1044 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1110

Qy       1104 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1170

Qy       1164 CTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCT 1223
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1171 CTAAAGTTTGGACTGCCCATCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCC 1230

Qy       1224 GCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1283
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 GTCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1290

Qy       1284 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1350

Qy       1344 TATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1403
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1351 TATTATGTGGTATCGACCGGGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1410

Qy       1404 GGAATGCCTATTCTTTTTGGTCTAAAA 1430
              |||||||||||||||||||||||||| |
Db       1411 GGAATGCCTATTCTTTTTGGTCTAAGA 1437
```

RESULT 4
US-10-425-115-18039
; Sequence 18039, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18039
; LENGTH: 1706
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116454C.1
US-10-425-115-18039

Query Match 82.3%; Score 1196.8; DB 18; Length 1706;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 1425; Conservative 0; Mismatches 7; Indels 214; Gaps 1;

```
Qy         24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 92

Qy         84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 152

Qy        144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 212

Qy        204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 272

Qy        264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 332

Qy        324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 392

Qy        384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCT----- 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGGTGA 452

Qy        439 ---------------------------------------------------------- 438

Db        453 GAAAAATTAACTTCAGCAGTGTGTGTGTATCTGGTGGTGGTGTTGGGGTCTTGGGTATGC 512

Qy        439 ---------------------------------------------------------- 438

Db        513 AATGTGGTGCATGCTATTTTCTGACGATATGAGTCCACTTAGCTGCCTTTCCTGAGAAAT 572

Qy        439 ---------------------------------------------------------- 438

Db        573 GTTTGTAGCAGCCTTGTAGACTGGGGGATGATTATGTTTCAGCTTAATCATCTCATACTA 632

Qy        439 ----------------------------GAGGAATCTAGCCAGGCACCAGTAAAGAATG 469
                                          ||||||||||||||||||||||||||||||||
Db        633 CCTGTCATGTTATTTTGTTTATGTTTCCAGAGGAATCTAGCCAGGCACCAGTAAAGAATG 692

Qy        470 TTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATA 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATA 752

Qy        530 CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTG 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTG 812

Qy        590 CCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTG 649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 CCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTG 872

Qy        650 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATT 709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATT 932

Qy        710 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 992

Qy        770 CCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 CCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 1052

Qy        830 GTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 889
```

```
Db       1053 GTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 1112

Qy        890 CTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 949

Db       1113 CTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 1172

Qy        950 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 1009

Db       1173 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 1232

Qy       1010 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAA 1069

Db       1233 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAA 1292

Qy       1070 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCC 1129

Db       1293 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCC 1352

Qy       1130 CCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTG 1189

Db       1353 CCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTG 1412

Qy       1190 TAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGC 1249

Db       1413 TAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGC 1472

Qy       1250 TTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGC 1309

Db       1473 TTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGC 1532

Qy       1310 TACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCA 1369

Db       1533 TACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCA 1592

Qy       1370 TCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCTAAA 1429

Db       1593 TCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCAAAA 1652

Qy       1430 AAAAAAAAAAAAAAAAAAAAAAAAA 1455

Db       1653 AAAAAAAAAAAAAAAGCGGCCAAAAA 1678


RESULT 5
US-10-425-114-24871
; Sequence 24871, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24871
;   LENGTH: 1649
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3689-210-C7_FLI
US-10-425-114-24871

  Query Match            81.0%;  Score 1179;  DB 17;  Length 1649;
  Best Local Similarity  86.8%;  Pred. No. 0;
  Matches 1403;  Conservative    0;  Mismatches    0;  Indels  214;  Gaps    1;

Qy         24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83

Db         33 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 92

Qy         84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143

Db         93 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 152

Qy        144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203

Db        153 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 212

Qy        204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263

Db        213 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 272

Qy        264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323

Db        273 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 332

Qy        324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383

Db        333 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 392

Qy        384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCT----- 438

Db        393 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGGTGA 452

Qy        439 ---------------------------------------------------------- 438

Db        453 GAAAAATTAACTTCAGCAGTGTGTGTGTATCTGGTGGTGGTGTTGGGGTCTTGGGTATGC 512

Qy        439 ---------------------------------------------------------- 438

Db        513 AATGTGGTGCATGCTATTTTCTGACGATATGAGTCCACTTAGCTGCCTTTCCTGAGAAAT 572

Qy        439 ---------------------------------------------------------- 438

Db        573 GTTTGTAGCAGCCTTGTAGACTGGGGGATGATTATGTTTCAGCTTAATCATCTCATACTA 632

Qy        439 ----------------------------GAGGAATCTAGCCAGGCACCAGTAAAGAATG 469

Db        633 CCTGTCATGTTATTTTGTTTATGTTTCCAGAGGAATCTAGCCAGGCACCAGTAAAGAATG 692

Qy        470 TTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATA 529

Db        693 TTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATA 752

Qy        530 CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTG 589

Db        753 CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTG 812

Qy        590 CCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTG 649

Db        813 CCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTG 872

Qy        650 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATT 709

Db        873 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATT 932

Qy        710 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 769

Db        933 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 992

Qy        770 CCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 829

Db        993 CCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 1052

Qy        830 GTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 889

Db       1053 GTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 1112

Qy        890 CTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 949
```

```
Db       1113 CTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 1172

Qy        950 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 1009

Db       1173 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 1232

Qy       1010 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAA 1069

Db       1233 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAA 1292

Qy       1070 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCC 1129

Db       1293 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCC 1352

Qy       1130 CCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTG 1189

Db       1353 CCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTG 1412

Qy       1190 TAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGC 1249

Db       1413 TAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGC 1472

Qy       1250 TTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGC 1309

Db       1473 TTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGC 1532

Qy       1310 TACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCA 1369

Db       1533 TACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCA 1592

Qy       1370 TCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCT 1426

Db       1593 TCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCT 1649


RESULT 6
US-10-767-701-13951
; Sequence 13951, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13951
;   LENGTH: 1537
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS2718_1
US-10-767-701-13951

  Query Match             76.5%;  Score 1113.4;  DB 18;  Length 1537;
  Best Local Similarity   90.9%;  Pred. No. 3.5e-304;
  Matches 1270;  Conservative    0;  Mismatches  101;  Indels   26;  Gaps    7;

Qy         51 GGCGCAGACGCCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGC 110

Db         63 GTCGTCGCCCGCAGACGCCCGGGGGGCCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGC 122

Qy        111 CAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCAT 170

Db        123 CAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCGCCGAGCTCTACACCCAGGCCAT 182

Qy        171 CGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCT 230

Db        183 CGACGCGGGGCCCGCCACCGCCGAACTCTACGCCGACCGCGCCCAGGCGCACATCAAGCT 242

Qy        231 CGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGAT 290

Db        243 CGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAACTTGATCCTACGAT 302

Qy        291 GCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGC 350

Db        303 GCATAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGC 362

Qy        351 AAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCT 410

Db        363 AAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTTACTCGTCT 422

Qy        411 ATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGT 470

Db        423 ATTGAAAGAATGTGAAGAGCGCATTGCTGAGGAATCTAGCCAGGCTCCAGCAAAGAATGT 482

Qy        471 TGAGGCTACTGTG---------GCTGCTACTATTGAGGACAAGGAGGATTTCACAAATAT 521

Db        483 TGAGCCTCCTGTGGCTGCTGCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATAT 542

Qy        522 GGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTA 581

Db        543 GGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATAGTCCCATGACTACTA 602

Qy        582 CAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGT 641

Db        603 CAACAGTGCCACAGAAGTGGTTTTGACAATATTTGCTAAGGGTGTTCCCGCTGATAGTGT 662

Qy        642 AGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACC 701

Db        663 AGTCATTGATTTTGGTGAACAGATGTTAAGTGTGTCCATTGAAGTCCCTGGTGAAGAACC 722

Qy        702 ATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGT 761

Db        723 ATACCATTTTCAGCCCCGTCTGTTTGCTAAGATTATCCCTGAGAAATGCAAATATCAAGT 782

Qy        762 CTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCT 821

Db        783 CTTGTCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCT 842

Qy        822 GGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGC 881

Db        843 GGATTATAGCGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGC 902

Qy        882 CCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGA 941

Db        903 CCCAAGACCTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGA 962

Qy        942 AGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTT 1001

Db        963 AGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTT 1022

Qy       1002 CCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGT 1061

Db       1023 CCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGT 1082

Qy       1062 GGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGA 1121

Db       1083 GGAATCTAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGAAGGTGGA 1142

Qy       1122 AGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGAC-TGCC 1180

Db       1143 AGGGAGCCCTCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTCGGACTTGCC 1202

Qy       1181 CGTCTTTTGTAAATCCAGGTC-TTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCC 1239

Db       1203 CGTCTTTTGTAAATCCAGGTCTTTGGAAATTATGATCTAGTAACGTCCACCCATTTGGGT 1262

Qy       1240 ATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGT 1299

Db       1263 GCTGAGCTTGGTTAAAGTCTTCTGCTTTTTGTAAGCTTTCTGTATGACAGTCTAGC--GT 1320
```

```
Qy       1300 GTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATA----GGTATGCTATTATGTGGTA 1355
             |||||||||||| |||||||||| ||||||||||||||    |||||||||| ||||||
Db       1321 GTCTGAGTGCTGCTTGCCGTAGTTTTTGATGGACATATCATGGTATGCTATTCTGTGGTA 1380

Qy       1356 TCGACCCCGAACCATCTGTGGTATCAAAGACAG-CAAACTATAATGTATGGAATGCCTAT 1414
             ||        | |||||||||| ||||||||  |||||||||| | ||||||||||||
Db       1381 TC--------AACATCTGTGGTAACAAAGACAGTAAAACTATAATCTGTGGAATGCCTAT 1432

Qy       1415 TCTTTTTGGTCTAAAAA 1431
             ||||||||||||||| |
Db       1433 TCTTTTTGGTCTAAAGA 1449


RESULT 7
US-10-425-114-25845
; Sequence 25845, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25845
;   LENGTH: 1361
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3960-013-B4_FLI
US-10-425-114-25845

  Query Match             72.0%;  Score 1047;  DB 17;  Length 1361;
  Best Local Similarity   92.3%;  Pred. No. 2.1e-285;
  Matches 1125;  Conservative    0;  Mismatches   90;  Indels    4;  Gaps    2;

Qy         54 GCAGACGCCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAA 113
              |||||||||| ||||||| || |||||||||||||||||||||||||||||| |||||
Db         71 GCAGACGCCCAGAGCCGGAGGGAGCGCCATGGCCGCGTCGGATCTGGAGAGCAAAGCCAA 130

Qy        114 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGA 173
              |||||||||||||||||||||||||||||||| |||| |||||||  ||||||||||
Db        131 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCGCCGAACTCTACACGCAGGCCATCGA 190

Qy        174 CGCCGGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 233
              ||||||||||||||| ||||| ||||| ||||||||||||||||||||||||||||||
Db        191 CGCCGGGCCCGCCACTGCCGATCTCTACGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 250

Qy        234 CAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCA 293
              |||||||||||||||||| |||||||||||||||||  |||||||||  |||||
Db        251 CAACTACACTGAGGCTGTGGCAGATGCTAACAAAGCAATTGGTCTTGATCCTACAATGCA 310

Qy        294 TAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 353
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 370

Qy        354 GGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATT 413
              ||||||||||||||| |||||||| |||||| |||||||||||||||| | ||||||||
Db        371 GGCTGCTCTTGAGTTAGGTTCTTCCTATGCACCAGGCGATTCAAGGTTTACCCGTCTATT 430

Qy        414 GAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGA 473
              |||||||||||||||| |||| |||||||||||||||||||||||| |||||||| ||
Db        431 GAAGGAATGTGATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGAATGTCGA 490

Qy        474 GGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACC 533
              |||| |||| |||||||||| |||||||||||||||||| || |||||||||| ||||||||
Db        491 GGCTCCTGTAGCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATATGGATAATACACC 550

Qy        534 ACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCAC 593
              |||||| |||||||||||||||||| ||||| ||||||||||||||||||||||||||
Db        551 GCCAGTGGTAGAACCCCCAAGCAAACCTAAATATAGGCATGACTACTACAACAGTGCCAC 610

Qy        594 AGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 653
              ||||||||| || |||||||| ||||||||||||||||||||||||||||||||||||||
Db        611 AGAAGTGGTCCTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 670

Qy        654 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCA 713
              |||||| |||||||||||||||||||||||| |||||||||||||| ||||||||||
Db        671 TGGTGATCAGATGTTGAGTGTATCCATTGAAGTACCTGGTGAAGAACCGTACCATTTTCA 730

Qy        714 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAA 773
              |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db        731 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAGTATCAAGTTTTATCCACCAA 790

Qy        774 GGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 833
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 GGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 850

Qy        834 AAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 893
              ||||||||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db        851 AAGACCAAAGGCTATTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 910

Qy        894 ATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGA 953
              |||||||||||||||| | ||||| |||||||||||||||||||||||| ||||||||
Db        911 ATACCCATCTTCAAAGTCAAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGA 970

Qy        954 GGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTA 1013
              |||||||||||||||||||||| |||||||||||||||||||| |||||||||| |||||
Db        971 GGAGAAGGAAGAAAAACTTGAAGGTGATGCTGCATTGAACAAGTTCTTCCGTGATATCTA 1030

Qy       1014 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGG 1073
              |||||||||||||||||||||||||||||||||    |||||||||  |||||| |||||
Db       1031 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGGACAAGTCATTCAGGGAATCTAATGG 1090

Qy       1074 CACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCC 1133
              || |||||||||||||||||||||||||||||| |||||| ||| |||| |||||| ||
Db       1091 TACCGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGACGGTGGAAGCGAGCCCTCC 1150

Qy       1134 TGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAA 1193
              |||||||||||||||||||||||||||||||    | ||||||||||||||||  || |||||||
Db       1151 TGATGGTATGGAGCTCAAGAAGTGGGAAATTTGAAGTTTGGACTGCCC--CTCTTGTAAA 1208

Qy       1194 TCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCA--TAGTGCCATGAGCTTGCTT 1251
              |||||||||||||  ||||||||| ||||| |||||||  | ||| | | || ||| ||
Db       1209 TCCAGGTCTTGGAAGTTATGACCTAGTTCTGTCCACCCATTTGGTGTCGTAAGTTTGGTT 1268

Qy       1252 GGTTAAGTCTCTGCTTTTG 1270
                 |  || | || |||||
Db       1269 AAGTGTGTGTGTGTTTTTG 1287


RESULT 8
US-10-425-114-22598
; Sequence 22598, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22598
;   LENGTH: 1490
```

```
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3591-005-E5_FLI
US-10-425-114-22598

  Query Match             72.0%;  Score 1047;  DB 17;  Length 1490;
  Best Local Similarity   92.3%;  Pred. No. 2.2e-285;
  Matches 1125;  Conservative    0;  Mismatches   90;  Indels    4;  Gaps    2;

Qy         54 GCAGACGCCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAA 113
              |||||||||||| ||||||| || |||||||||||||||||||||||||||||||| |||||
Db        139 GCAGACGCCCAGAGCCGGAGGGAGCGCCATGGCCGCGTCGGATCTGGAGAGCAAAGCCAA 198

Qy        114 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGA 173
              |||||||||||||||||||||||||||||||||||| |||| |||||||  ||||||||||
Db        199 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCGCCGAACTCTACACGCAGGCCATCGA 258

Qy        174 CGCCGGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 233
              |||||||||||||| ||||| ||||| ||||||||||||||||||||||||||||||||
Db        259 CGCCGGGCCCGCCACTGCCGATCTCTACGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 318

Qy        234 CAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCA 293
              |||||||||||||||||||| |||||||||||||||||||  ||||||||||  |||||
Db        319 CAACTACACTGAGGCTGTGGCAGATGCTAACAAAGCAATTGGTCTTGATCCTACAATGCA 378

Qy        294 TAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 353
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 438

Qy        354 GGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATT 413
              |||||||||||||| |||||||| |||||| |||||||||||||||| | ||||||||
Db        439 GGCTGCTCTTGAGTTAGGTTCTTCCTATGCACCAGGCGATTCAAGGTTTACCCGTCTATT 498

Qy        414 GAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGA 473
              |||||||||||||||| |||| |||||||||||||||||||||||| ||||||||| ||
Db        499 GAAGGAATGTGATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGAATGTCGA 558

Qy        474 GGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACC 533
              |||| |||| ||||||||| |||||||||||||||| || |||||||||  ||||||||
Db        559 GGCTCCTGTAGCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATATGGATAATACACC 618

Qy        534 ACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCAC 593
              |||||| |||||||||||||||||| ||||| ||||||||||||||||||||||||||
Db        619 GCCAGTGGTAGAACCCCCAAGCAAACCTAAATATAGGCATGACTACTACAACAGTGCCAC 678

Qy        594 AGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 653
              ||||||||| || ||||||| ||||||||||||||||||||||||||||||||||||||
Db        679 AGAAGTGGTCCTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 738

Qy        654 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCA 713
              |||||| |||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db        739 TGGTGATCAGATGTTGAGTGTATCCATTGAAGTACCTGGTGAAGAACCGTACCATTTTCA 798

Qy        714 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAA 773
              |||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db        799 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAGTATCAAGTTTTATCCACCAA 858

Qy        774 GGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 833
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 918

Qy        834 AAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 893
              |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db        919 AAGACCAAAGGCTATTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 978

Qy        894 ATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGA 953
              |||||||||||||||| | ||||| |||||||||||||||||||||||| ||||||||
Db        979 ATACCCATCTTCAAAGTCAAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGA 1038

Qy        954 GGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTA 1013
              |||||||||||||||||||| |||||||||||||||||||| |||||||||| |||||
Db       1039 GGAGAAGGAAGAAAAACTTGAAGGTGATGCTGCATTGAACAAGTTCTTCCGTGATATCTA 1098

Qy       1014 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGG 1073
              |||||||||||||||||||||||||||||||||||||||    |||||||||  |||||| |||||
Db       1099 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGGACAAGTCATTCAGGGAATCTAATGG 1158

Qy       1074 CACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCC 1133
              || |||||||||||||||||||||||||||||||||||||| |||||| ||| |||| |||||| ||
Db       1159 TACCGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGACGGTGGAAGCGAGCCCTCC 1218

Qy       1134 TGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAA 1193
              |||||||||||||||||||||||||||||    | ||||||||||||||  || |||||||
Db       1219 TGATGGTATGGAGCTCAAGAAGTGGGAAATTTGAAGTTTGGACTGCCC--CTCTTGTAAA 1276

Qy       1194 TCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCA--TAGTGCCATGAGCTTGCTT 1251
              |||||||||||||| ||||||||| ||||| ||||||| | ||| | | || ||| ||
Db       1277 TCCAGGTCTTGGAAGTTATGACCTAGTTCTGTCCACCCATTTGGTGTCGTAAGTTTGGTT 1336

Qy       1252 GGTTAAGTCTCTGCTTTTG 1270
                 |  || | || |||||
Db       1337 AAGTGTGTGTGTGTTTTTG 1355


RESULT 9
US-10-425-115-18041
; Sequence 18041, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18041
;   LENGTH: 1519
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_116456C.1
US-10-425-115-18041

  Query Match             71.9%;  Score 1046;  DB 18;  Length 1519;
  Best Local Similarity   92.3%;  Pred. No. 4.3e-285;
  Matches 1124;  Conservative    0;  Mismatches   90;  Indels    4;  Gaps    2;

Qy         54 GCAGACGCCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAA 113
              |||||||||||| ||||||| || |||||||||||||||||||||||||||||||| |||||
Db        139 GCAGACGCCCAGAGCCGGAGGGAGCGCCATGGCCGCGTCGGATCTGGAGAGCAAAGCCAA 198

Qy        114 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGA 173
              |||||||||||||||||||||||||||||||||||| |||| |||||||  ||||||||||
Db        199 GGAGGCCTTCGTCGACGACGACTTCGAGCTGGCCGCCGAACTCTACACGCAGGCCATCGA 258

Qy        174 CGCCGGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 233
              |||||||||||||| ||||| ||||| ||||||||||||||||||||||||||||||||
Db        259 CGCCGGGCCCGCCACTGCCGATCTCTACGCCGACCGCGCCCAGGCGCACATCAAGCTCGG 318

Qy        234 CAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCA 293
              |||||||||||||||||||| |||||||||||||||||||  ||||||||||  |||||
Db        319 CAACTACACTGAGGCTGTGGCAGATGCTAACAAAGCAATTGGTCTTGATCCTACAATGCA 378

Qy        294 TAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 353
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAA 438

Qy        354 GGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATT 413
              |||||||||||||| |||||||| |||||| |||||||||||||||| | ||||||||
Db        439 GGCTGCTCTTGAGTTAGGTTCTTCCTATGCACCAGGCGATTCAAGGTTTACCCGTCTATT 498
```

```
Qy        414 GAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGA 473
Db        499 GAAGGAATGTGATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGAATGTCGA 558

Qy        474 GGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACC 533
Db        559 GGCTCCTGTAGCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATATGGATAATACACC 618

Qy        534 ACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCAC 593
Db        619 GCCAGTGGTAGAACCCCCAAGCAAACCTAAATATAGGCATGACTACTACAACAGTGCCAC 678

Qy        594 AGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 653
Db        679 AGAAGTGGTCCTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTT 738

Qy        654 TGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCA 713
Db        739 TGGTGATCAGATGTTGAGTGTATCCATTGAAGTACCTGGTGAAGAACCGTACCATTTTCA 798

Qy        714 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAA 773
Db        799 GCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAGTATCAAGTTTTATCCACCAA 858

Qy        774 GGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 833
Db        859 GGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGG 918

Qy        834 AAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 893
Db        919 AAGACCAAAGGCTATTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTC 978

Qy        894 ATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGA 953
Db        979 ATACCCATCTTCAAAGTCAAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGA 1038

Qy        954 GGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTA 1013
Db       1039 GGAGAAGGAAGAAAAACTTGAAGGTGATGCTGCATTGAACAAGTTCTTCCGTGATATCTA 1098

Qy       1014 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGG 1073
Db       1099 CAAGGATGCTGATGAAGATATGCGGAGGGCCATGGACAAGTCATTCAGGGAATCTAATGG 1158

Qy       1074 CACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCC 1133
Db       1159 TACCGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGACGGTGGAAGCGAGCCCTCC 1218

Qy       1134 TGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAA 1193
Db       1219 TGATGGTATGGAGCTCAAGAAGTGGGAAATTTGAAGTTTGGACTGCCC--CTCTTGTAAA 1276

Qy       1194 TCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCA--TAGTGCCATGAGCTTGCTT 1251
Db       1277 TCCAGGTCTTGGAAGTTATGACCTAGTTCTGTCCACCCATTTGGTGTCGTAAGTTTGGTT 1336

Qy       1252 GGTTAAGTCTCTGCTTTT 1269
Db       1337 AAGTGTGTGTGTGTTTTT 1354
```

RESULT 10
US-10-739-930-4858/c
; Sequence 4858, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 4858
; LENGTH: 2471
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2471)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER396_1
US-10-739-930-4858

```
  Query Match             51.8%;  Score 753;  DB 18;  Length 2471;
  Best Local Similarity   79.3%;  Pred. No. 5.6e-202;
  Matches  938;  Conservative    0;  Mismatches  200;  Indels   45;  Gaps    2;

Qy         61 CCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 120
Db       2303 CTCGACGCAGACATGGCCGCCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 2244

Qy        121 TTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGG 180
Db       2243 TTCGTCGACGACGACTTCGAGCTGGCCGCCGAGCTCTACACCCAGGCCATTGAGGCGGGG 2184

Qy        181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240
Db       2183 CCGGCCACCGCGGAACTCTACGCCGACCGAGCCCAGGCGCACATCAAGCTGGGCAGTTAC 2124

Qy        241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300
Db       2123 ACTGAGGCTGTAGCTGATGCCAACAAAGCAATTGAACTTGATCCTTCGATGCATAAAGCA 2064

Qy        301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360
Db       2063 TACCTTCGGAAGGGCTCTGCTTGCATCAAGCTGGAGGAATACCAAACTGCAAAGGCTGCT 2004

Qy        361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420
Db       2003 CTTGAAGTGGGTTCTTCTTATGCATCTGGTGACTCGAGGTTTACTCGTCTGATGAAGGAG 1944

Qy        421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480
Db       1943 TGTGATGATCGTATTGCTGAGGAGGCTAGCCAGGTGCCAGTAAAGAATGCCGCTGCGGCT 1884

Qy        481 GTGGCT------------------------------------------GCTACTATTGAG 498
Db       1883 GTTGCTTCAGCTACATCTTCGGGGGCATCTTCCGGGGCTACAACTGTGGCTACTGAAGCT 1824

Qy        499 GACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAA 558
Db       1823 GAGGACCAGGATGGTGCAAATATGGAGAATGCACAGCCAACGATAGAAGTGCCAAGCAAG 1764

Qy        559 CCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCT 618
Db       1763 CCCAAATACAGGCATGACTACTACAATACTCCTACAGAAGTGGTACTGACTATATTTGCT 1704

Qy        619 AAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCC 678
Db       1703 AAGGGTGTTCCAGCTGACAGCGTGGTTGTTGACTTTGGTGAACAGATGCTGAGTGTCTCA 1644

Qy        679 ATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATC 738
Db       1643 ATTGAACTTCCCGGTGAGGAACCATACCATTTTCAGCCTCGTCTGTTTTCAAAGATCGTC 1584

Qy        739 CCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCT 798
Db       1583 CCAGATAAGTGCAAGTATACTGTGTTGTCTACAAAGGTTGAAATGCGCCTTGCAAAAGCT 1524

Qy        799 GAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAG 858
Db       1523 GAGCCAGTAACTTGGACATCATTGGATTATACTGGTAAACCAAAG---GCTCCTCAGAAG 1467

Qy        859 ATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAA 918
```

```
Db       1466 ATAAATGTACCAGCTGAATCAGCCCAGAGGCCATCTTATCCATCATCAAAATCCAAAAAG 1407

Qy        919 GACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGT 978

Db       1406 GACTGGGATAAGCTTGAAGCTGAAGTGAAAAAACAGGAGAAGGATGAAAAACTTGATGGC 1347

Qy        979 GATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGG 1038

Db       1346 GATGCTGCATTGAACAAATTTTTCCGTGAAATTTACAGTGATGCTGATGAAGATATGCGT 1287

Qy       1039 AGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAA 1098

Db       1286 AGAGCAATGATGAAGTCTTTTGTGGAGTCTAATGGAACCGTTCTCTCAACCAACTGGAAG 1227

Qy       1099 GATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGG 1158

Db       1226 GATGTCGGGAAAAAGACGGTTGAAGGAAGCCCTCCTGATGGAATGGAGCTCAAGAAGTGG 1167

Qy       1159 GAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTC 1201

Db       1166 GAGTATTAATGAATCCAGATCATCCATGTTGTAAATCCTGGTC 1124


RESULT 11
US-10-437-963-37229
; Sequence 37229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37229
;   LENGTH: 1569
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_40977C.1
US-10-437-963-37229

  Query Match             48.4%;  Score 704.8;  DB 18;  Length 1569;
  Best Local Similarity   78.3%;  Pred. No. 2e-188;
  Matches  904;  Conservative    0;  Mismatches  202;  Indels   48;  Gaps    3;

Qy          63 CAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTT 122

Db          72 CCAAGCCGCCATGGCAACCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTT 131

Qy         123 CGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCC 182

Db         132 CGTCGACGACGACTTCGAGCTCGCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCC 191

Qy         183 CGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACAC 242

Db         192 CGCCACCGCCGAGCTCTACGCCGACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACAC 251

Qy         243 TGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTA 302

Db         252 TGAGGCTGTAGCTGATGCTAACAAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTA 311

Qy         303 CTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCT 362

Db         312 TCTTCGTAAAGGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCT 371

Qy         363 TGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATG 422

Db         372 TGAATTGGGTTACTCGTTCGCATCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTG 431

Qy         423 TGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGG------- 475

Db         432 TGATGAGCGCATTGCTGAGGAGCTTAGTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGC 491

Qy         476 --------CTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 527

Db         492 AGCTGCCCCCTCTGTTGCTTCTTTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAA 551

Qy         528 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 587

Db         552 TACACCACCAATGGTAGA------AGTGAAGCCAAAATACAGGCACGACTTCTACAACAG 605

Qy         588 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 647

Db         606 TGCTACAGAAGTTGTATTGACAATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGT 665

Qy         648 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 707

Db         666 TGATTTTGGTGAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCA 725

Qy         708 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 767

Db         726 TTTTCAGCCTCGTCTGTTTTCTAAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATC 785

Qy         768 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 827

Db         786 CACGAAGGTTGAAATAAGACTGGCTAAAGCTGAACAGATTACATGGACCTCACTTGATTA 845

Qy         828 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGC------------------- 868

Db         846 TGATAAAAAACCAAAGGCTGTTCCACAAAAGATAATCCCTCCAGTTTTACTGATAGCCCT 905

Qy         869 --------CAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGA 920

Db         906 TCTTTTATCAGCTGAATCGGCCCAGAGGCCATCATATCCTTCCTCAAAATCCAAGAAAGA 965

Qy         921 CTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGA 980

Db         966 CTGGGATAAACTGGAAGCTGAAGTTAAAAAGGAGGAGAAGGAGGAGAAGCTTGAAGGCGA 1025

Qy         981 TGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAG 1040

Db        1026 TGCTGCATTGAACAAATTTTTCCGTGACATCTACAGTGATGCTGATGAAGACATGCGACG 1085

Qy        1041 GGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGA 1100

Db        1086 AGCAATGATGAAATCTTTTGTTGAATCTAACGGTACTGTTCTGTCGACCAATTGGAAAGA 1145

Qy        1101 TGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGA 1160

Db        1146 TGTTGGCTCGAAGAAGGTAGAGGGAAGCCCACCTGATGGGATGGAGCTTAAGAAATGGGA 1205

Qy        1161 ATACTAAAGTTTGG 1174

Db        1206 GTACTAAGATCTCG 1219


RESULT 12
US-10-074-473-2
; Sequence 2, Application US/10074473
; Publication No. US20030167504A1
; GENERAL INFORMATION:
;  APPLICANT: Crane, Edmund H. III
;  TITLE OF INVENTION: Maize Rar1 Interactor Polynucleotides
;  TITLE OF INVENTION:  and Methods of Use
;  FILE REFERENCE: 35718/241314
;  CURRENT APPLICATION NUMBER: US/10/074,473
```

```
;  CURRENT FILING DATE:  2002-02-12
;  PRIOR APPLICATION NUMBER: 60/268,157
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 571
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-074-473-2

  Query Match             39.2%;  Score 571;  DB 16;  Length 571;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-151;
  Matches  571;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAATTCGGCACGAGCGGCACGAGAGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACG 60
Db          1 GAATTCGGCACGAGCGGCACGAGAGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACG 60

Qy         61 CCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 120
Db         61 CCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 120

Qy        121 TTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGG 180
Db        121 TTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGG 180

Qy        181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240
Db        181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240

Qy        241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300
Db        241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300

Qy        301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360
Db        301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360

Qy        361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420
Db        361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420

Qy        421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480
Db        421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480

Qy        481 GTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTG 540
Db        481 GTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTG 540

Qy        541 ATAGAACCCCCAAGCAAACCAAAATACAGGC 571
Db        541 ATAGAACCCCCAAGCAAACCAAAATACAGGC 571

RESULT 13
US-10-424-599-65699
; Sequence 65699, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65699
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_30341C.1
US-10-424-599-65699

  Query Match             32.9%;  Score 479;  DB 17;  Length 1691;
  Best Local Similarity   64.7%;  Pred. No. 1.9e-124;
  Matches  729;  Conservative    0;  Mismatches  395;  Indels    3;  Gaps    1;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db        178 GCTTCCGATCTCGAAGCAAAAGCCAAAGAGGCCTTCGTCGAAGACCACTTCGAACTCGCC 237

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db        238 GTGGACCTTCTCTCTCAGGCCATTCATCTCGAACCCAACAAAGCCGAATTCTACGCCGAC 297

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        298 CGCGCCCAAGCCAACATCAAACTCAACAACTTCACTGAGGCTGTGGCTGATGCTAACAAA 357

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        358 GCAATTGAATTGAATCCTTCTCTACCAAAAGCATATTTGCGGAAAGGTACTGCTTGCATG 417

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        418 AAGCTTGAGGAATATGAAACTGCAAAGGCTGCTCTGGAGGTGGGTGCATCATTGTCTCCA 477

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        478 GACAACTCAAGGTTTGCTACTTTGATCAAAGAATGTGATAAGCTCATTGCAGAAGAATCT 537

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        538 TATACCCCACCTATTATAGAGGAAAAGGCCACAACAGAAGATGTTACTCCA---AAAGAT 594

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        595 GTTCAGCAACAAGATGATCTTCTAGAGAAGCCAACAGTTACAGTGACTAAACCTAAATAT 654

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        655 AGGCATGAATTCTACCAGAAACATGATCAGGTGGTTGTGACCATATTTGCAAAGAAAATC 714

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        715 CCAAAGGAAAGCATTACTGTTGACTTCGGTGAACAAATACTAAGTGTTAGTATTAATGTC 774

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        775 ACAGGAGAAGACACATATGTTTTTCAACCTCGCTTATTTGGAAAGATCATACCTTCCAAT 834

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        835 TGTCAATATGAAGTTTTGTCCACCAAAATTGAAATTCGTCTTGCAAAAGCAGAACCTATT 894

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        895 CACTGGACATCTCTAGAATTCACCAGAGATAGCGTTGTTCAACAGAGGGTTAGTGCTTTG 954

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        955 TCAGTTACTGGAGGTGAAAACCTACTTATCCATCCTCAAAACGCTGCGAAAATACGACA 1014

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db       1015 GAAGGGGAAGCTCAAGTCAAGAAAGAGGGGAAAGATGAAAAGCTTGATGGCGATGCTGCA 1074

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db       1075 TTGAACAAATTTTTCCGCGAAATATATCAAGATGCAGATGAGGATACAAGAAGAGCAATG 1134
```

```
Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db       1135 AAAAAGTCATTTGTTGAGTCTAATGGGACAGTTCTGTCTACAAACTGGAAAGAAGTGGGA 1194

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
Db       1195 ACAAAGAAGGTAGAGGGAAGTCCTCCTGATGGCATGGAGTTGAAGAAATGGGAATATTAA 1254

Qy       1168 AGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGA 1214
Db       1255 GACATTTTCTGGATCAGTGTTGCATATATTTTTATGTGGGATTATTA 1301
```

RESULT 14
US-10-424-599-81927
; Sequence 81927, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 81927
;   LENGTH: 1598
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_44998C.1
US-10-424-599-81927

```
  Query Match             32.7%;  Score 476.2;  DB 17;  Length 1598;
  Best Local Similarity   66.2%;  Pred. No. 1.1e-123;
  Matches  721;  Conservative    0;  Mismatches  358;  Indels   10;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db        145 GCTTCCGATCTTGAGCTTAAGGCCAAAGAGGCTTTCGAAGATGATAACTATGATCTCGCC 204

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db        205 TACGACCTCTTAACTCAGGCCATTGGTCTCAGCCCCAACAACGCAGACCTATATGCTGAC 264

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        265 CGTGCGCAAGTCAACATCAAAGTCAACAACCTCACCGAGGCTGTTTCTGATGCAAACAAG 324

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        325 GCGATTGAGTTGAATCCTTCTCACTCAAAAGCATATTTGCGAAAAGGTACCGCATGCATC 384

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        385 AAGCTTGAGGAATATCAGACTGCTAAGGCAGCTCTAGAGATGGGTGCTTCATTGGCTCCT 444

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        445 GGAGATTCTAAATTTACTGATTTGATCAAAGACTGCGATGAACTGATTGCAGAAGAATCT 504

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        505 GGTGTCATACCCATACAAGAAGAGAGCACAACACAGGGTGCTGCTACAAAAGCTGTTGAG 564

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        565 G------CAGAGAATGATCTTCCAGAGCCACCTACAGTAACCGTGGTTAAACCTAAATAC 618

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        619 AGGCATGAATTCTACCAGAAACCTGATGAAATGGTTATTACCATATTTGCAAAGGGCATT 678

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        679 CCACGAGACAGCATTACTGTTGACTTTGGTGAACAAATATTAAGTGTTACAATTAATATC 738

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        739 CCTTGCAAAGATGCTTATGTCTTCCAACCTCGCTTATTTGGAAAGATCATACCCTCCAAA 798

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        799 TGCCGGTATGAAGTTTTGTCCACCAAAATTGAAATTTGCCTTGCAAAAGCAGATCATATC 858

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        859 CAATGGACATCTCTAGAATTCAACAAGGGTAGCACAGTTGCACAGAGGTTTAGTGTTTTG 918

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAA----AGACTG 923
Db        919 CCAGTTGCAAGAGGTGAAAAACCTACTTATCCATCCTCAAAACCGAAAAAATACAGATTG 978

Qy        924 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 983
Db        979 GGATAAGCTTGAAGCTCAAGTTAAGAAAGAGGAGAAAGAAGAAAAACTTGATGGTGATGC 1038

Qy        984 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1043
Db       1039 TGCGTTGAACAAATTTTTCCGTGATATATATCAAGATGCAGATGAGGACACAAGAAGAGC 1098

Qy       1044 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1103
Db       1099 AATGAGCAAATCATTTGTGGAGTCTAATGGAACAGTACTGTCTACAAACTGGAAAGAAGT 1158

Qy       1104 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1163
Db       1159 GGGATCAATGAAGGTACAGGAAAGTCCTCCTGATGGCATGGAGTTAAAGAAATGGTAATA 1218

Qy       1164 CTAAAGTTT 1172
Db       1219 TTGAGATTT 1227
```

RESULT 15
US-10-424-599-118531
; Sequence 118531, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 118531
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_78041C.1
US-10-424-599-118531

```
  Query Match             28.6%;  Score 416.6;  DB 17;  Length 483;
  Best Local Similarity   91.9%;  Pred. No. 4.3e-107;
  Matches  440;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        364 GAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGT 423
```

```
                 | ||  |||||||| ||||||  |||||||||||||||||  | ||||||||||||||||||
Db          5 GCGTCCGGTTCTTCCTATGCACCAGGCGATTCAAGGTTTACCCGTCTATTGAAGGAATGT 64

Qy        424 GATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTG 483
              ||||||  ||||  ||||||||||||||||||||||||  |||||||||  ||||||  ||||
Db         65 GATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGAATGTCGAGGCTCCTGTA 124

Qy        484 GCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATA 543
              |||||||||  ||||||||||||||||  ||  ||||||||||  ||||||||  ||||||  ||
Db        125 GCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATATGGATAATACACCGCCAGTGGTA 184

Qy        544 GAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTA 603
              ||||||||||||||||  |||||  |  ||||||||||||||||||||||||||||||||
Db        185 GAACCCCCAAGCAAACCTAAATATACGCATGACTACTACAACAGTGCCACAGAAGTGGTC 244

Qy        604 CTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAG 663
              ||  |||||||  ||||||||||||||||||||||||||||||||||||||  ||||||  |||
Db        245 CTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTGTGGTGATCAG 304

Qy        664 ATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTG 723
              ||||||||||||||||||||||  ||||||||||||  |||||||||  ||||||||||||
Db        305 ATGTTGAGTGTATCCATTGAAGTACCTGGTGAAGAACCGTACCATTTTTAGCCCCGTCTG 364

Qy        724 TTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATA 783
              ||  ||||||||||||||||||||||||  |||||  ||  ||||||||||||||  ||||||
Db        365 TTATCTAAGATTATCCCTGAGAAATGCAAGTATCACGTTCTATCCACCAAGGTCGAAATA 424

Qy        784 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAA 842
              |||||||||||||||||||||||||||||||||||  ||  ||||||||  ||||||||||||
Db        425 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACTCTTGATTATAGCGGAAGACCAAA 483
```

Search completed: April 23, 2005, 09:25:57
Job time : 922.465 secs

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 00:00:36 ; Search time 5229.64 Seconds
(without alignments)
10590.311 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcggcacgagcggcac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1378.2	94.7	1646	3	AY103953	AY103953 Zea mays
c 2	719	49.4	727	6	CA401716	CA401716 EL01N0424
3	708.2	48.7	752	6	CD443320	CD443320 EL01N0424
4	699	48.0	1131	9	CL959360	CL959360 OsIFCC002
c 5	660.6	45.4	686	6	CF041238	CF041238 QCI23f02.
6	657.2	45.2	772	6	CA258233	CA258233 SCCCRT300
7	652.4	44.8	668	2	AW076274	AW076274 614063D11
c 8	650.8	44.7	798	6	CA176210	CA176210 SCJLST102
9	647	44.5	756	6	CA184150	CA184150 SCRFST314
10	643.4	44.2	659	2	AW090936	AW090936 614068G07
11	641	44.1	759	6	CA245820	CA245820 SCEZFL508
12	637.4	43.8	778	6	CA227695	CA227695 SCJLFL301
13	631	43.4	666	6	CA215140	CA215140 SCSBAD112
14	628.8	43.2	681	2	AW076372	AW076372 614066C10
15	627	43.1	658	2	AW155786	AW155786 614096F11
c 16	623.2	42.8	648	1	AI603781	AI603781 486096G09
17	615.4	42.3	665	6	CA153735	CA153735 SCVPRZ203
18	608.2	41.8	678	6	CA177070	CA177070 SCMCST105
19	601.8	41.4	725	6	CA094787	CA094787 SCCCCL401
c 20	601	41.3	609	7	CK826852	CK826852 zmrsub1_0
c 21	600.2	41.3	609	2	AW017784	AW017784 614063D11
22	599.4	41.2	640	6	CF002911	CF002911 QBH17b05.
23	596.4	41.0	622	6	CA229980	CA229980 SCJFFL3C0
c 24	595.8	40.9	599	2	AW331768	AW331768 614096F11
25	594.8	40.9	730	2	BE362001	BE362001 DG1_83_H0
c 26	591	40.6	591	2	AW018202	AW018202 614068G07
c 27	589.8	40.5	593	4	BG265701	BG265701 1000034F0
c 28	584.4	40.2	594	1	AI601072	AI601072 486096G09
29	583.4	40.1	586	2	AW042392	AW042392 614028B03
30	583.4	40.1	609	6	CA229238	CA229238 SCAGFL302
31	579.8	39.8	718	6	CA106435	CA106435 SCQGHR101
32	576.8	39.6	702	6	CA262995	CA262995 SCPRLB202
33	576	39.6	588	5	BU499669	BU499669 946178B11
34	571.2	39.3	680	6	CA128103	CA128103 SCAGLR203
c 35	567.8	39.0	578	2	AW000264	AW000264 614009H02
36	565.2	38.8	594	6	CA098274	CA098274 SCMCCL605
37	563	38.7	574	4	BI644145	BI644145 949022C08
38	560.6	38.5	914	6	CA084222	CA084222 SCEQAM203
c 39	560.4	38.5	579	2	AW054458	AW054458 660007D01
c 40	560.2	38.5	607	1	AI601071	AI601071 486096G09
c 41	558.4	38.4	560	2	AW018065	AW018065 614066C10
42	557.2	38.3	630	6	CA486265	CA486265 WHE4329_D
43	556.4	38.2	578	6	CD996630	CD996630 QBC6a07.x
44	555.6	38.2	579	2	AW066518	AW066518 660015G08
c 45	555.6	38.2	736	6	CA227777	CA227777 SCJLFL301

ALIGNMENTS

RESULT 1
AY103953
LOCUS AY103953 1646 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO098564 mRNA sequence.
ACCESSION AY103953
VERSION AY103953.1 GI:21207031
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1646)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1646)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES Location/Qualifiers
source 1. .1646
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635643"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
ORIGIN

Query Match 94.7%; Score 1378.2; DB 3; Length 1646;

```
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1389;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy         24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        143 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAACCGGAAGGGGCGCCAT 202

Qy         84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 262

Qy        144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 322

Qy        204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 382

Qy        264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 442

Qy        324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 502

Qy        384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 562

Qy        444 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 622

Qy        504 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 682

Qy        564 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 742

Qy        624 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 802

Qy        684 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 743
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 AGTTCCTGGTGAAGAACCGTACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGA 862

Qy        744 GAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCA 803
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        863 GAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCA 922

Qy        804 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAG 863
              |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db        923 GGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTATTCCCCAGAAGATAAG 982

Qy        864 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTG 923
              |||||||||||||||||||||||||||||||||||||||||||||| | ||||| |||||
Db        983 CACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGTCAAAAAAGGACTG 1042

Qy        924 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 983
              |||||||||||||||||||| |||||||||||||||||||||||||||| ||||||||
Db       1043 GGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAAGAAAAACTTGAAGGTGATGC 1102

Qy        984 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1043
              |||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db       1103 TGCATTGAACAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 1162

Qy       1044 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1222

Qy       1104 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1282

Qy       1164 CTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCT 1223
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1283 CTAAAGTTTGGACTGCCCATCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCC 1342

Qy       1224 GCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1283
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1343 GTCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTA 1402

Qy       1284 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1403 TGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGC 1462

Qy       1344 TATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1403
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1463 TATTATGTGGTATCGACCGGGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTAT 1522

Qy       1404 GGAATGCCTATTCTTTTTGGTCTAAAA 1430
              |||||||||||||||||||||||| |
Db       1523 GGAATGCCTATTCTTTTTGGTCTAAGA 1549
```

RESULT 2
CA401716/c

```
LOCUS       CA401716                727 bp    mRNA    linear   EST 07-NOV-2002
DEFINITION  EL01N0424G10.g Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION   CA401716
VERSION     CA401716.1  GI:24766567
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 727)
  AUTHORS   Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F.,
            Larkins,B., Becraft,P. and Messing,J.
  TITLE     Characterization of the maize endosperm transcriptome and its
            comparison to the rice genome
  JOURNAL   Genome Res. 14 (10), 1932-1937 (2004)
COMMENT     Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .727
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm_4"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN
```

```
  Query Match             49.4%;  Score 719;  DB 6;  Length 727;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-175;
  Matches  722;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        712 CAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACC 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACC 668
```

```
Qy        772 AAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGT 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGT 608

Qy        832 GGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCT 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 GGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCT 548

Qy        892 TCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAG 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||.
Db        547 TCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAG 488

Qy        952 GAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATC 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATC 428

Qy       1012 TACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAAT 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 TACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAAT 368

Qy       1072 GGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 308

Qy       1132 CCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTA 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTA 248

Qy       1192 AATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGCTT 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTGCTT 188

Qy       1252 GGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGCTA 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGCTA 128

Qy       1312 CTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCATC 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACCATC 68

Qy       1372 TGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCTAAAAA 1431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||     |||||
Db          67 TGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTAAAAAAAAAA 8

Qy       1432 AAAAAAA 1438
              |||||||
Db           7 AAAAAAA 1
```

```
RESULT 3
CD443320
LOCUS       CD443320                752 bp    mRNA    linear   EST 03-JUN-2003
DEFINITION  EL01N0424G10.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION   CD443320
VERSION     CD443320.1  GI:31358963
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F.,
            Larkins,B., Becraft,P. and Messing,J.
  TITLE     Characterization of the maize endosperm transcriptome and its
            comparison to the rice genome
  JOURNAL   Genome Res. 14 (10), 1932-1937 (2004)
COMMENT     Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm_4"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN
```

```
  Query Match           48.7%;  Score 708.2;  DB 6;  Length 752;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-172;
  Matches  710;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         24 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 AGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGGCGCCAT 99

Qy         84 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 GGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCT 159

Qy        144 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGC 219

Qy        204 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAA 279

Qy        264 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATG 339

Qy        324 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGC 399

Qy        384 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 ATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGA 459

Qy        444 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 ATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAA 519

Qy        504 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAA 579

Qy        564 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 ATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGG 639

Qy        624 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 699

Qy        684 AGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTA 736
              |||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        700 AGTTCCTGGGTGAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTA 752
```

```
RESULT 4
CL959360
LOCUS       CL959360                1131 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC002721 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL959360
```

```
VERSION     CL959360.1  GI:52373436
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1131)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .1131
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

  Query Match          48.0%;  Score 699;  DB 9;  Length 1131;
  Best Local Similarity   78.6%;  Pred. No. 3.3e-170;
  Matches  894;  Conservative    0;  Mismatches  195;  Indels   48;  Gaps    3;

Qy         73 AGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGAC 132
              | ||   ||   |||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 ATGGCAACCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGAC 60

Qy        133 GACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCC 192
              |||||||||| ||| ||||||||||||  ||||| ||||| ||| | ||||||||||||
Db         61 GACTTCGAGCTCGCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCC 120

Qy        193 GACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTG 252
              || ||||| |||||||||||||||||| || |||||||| ||||||||||||||||||||
Db        121 GAGCTCTACGCCGACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTA 180

Qy        253 GCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAA 312
              || |||||||||| || ||||| ||||| ||     ||||| || |||||    || |||
Db        181 GCTGATGCTAACAAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAA 240

Qy        313 GGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGT 372
              || |||||||| ||    || || || || |||||||||| || |||||||| ||||||
Db        241 GGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCGGCTCTTGAATTGGGT 300

Qy        373 TCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGC 432
              |  || |  ||||| || || ||||||||| |||| ||| ||||||| ||||||||||||
Db        301 TACTCGTTCGCATCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGC 360

Qy        433 ATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGA---------------GGCT 477
              || ||||||||   |||  | |  || || ||||| | |||               | |
Db        361 ATTGCTGAGGAGCTTAGTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCGCCC 420

Qy        478 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 537
               |||| ||| ||  | |||||||| ||||| |||    |||| ||||| ||||||||||||
Db        421 TCTGTTGCTTCTTTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCA 480

Qy        538 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 597
               || ||||       |    || |||||||||||||| |||| ||||||||||| ||||||
Db        481 ATGGTAGA------AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAA 534

Qy        598 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 657
              || ||| ||||||| ||||| |||||||||||||||| | ||| ||  |||||||||||
Db        535 GTTGTATTGACAATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGT 594

Qy        658 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 717
              ||||| ||||| ||||| || |||||||| ||||| || || || ||||||||||||||
Db        595 GAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCT 654

Qy        718 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTT 777
              |||||||||||||||| |||||||||||| ||| ||| |||||  ||||||| ||||||
Db        655 CGTCTGTTTTCTAAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTT 714

Qy        778 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 837
              |||||| | || || |||||||| ||| | |||||||  | || ||||||  |  || |
Db        715 GAAATAAGACTGGCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAA 774

Qy        838 CCAAAGACTGTTCCCCAGAAGATAAGCACGC--------------------------CA 870
              |||||| ||||||| || ||||||| | | |                          ||
Db        775 CCAAAGGCTGTTCCACAAAAGATAATCCCTCCAGTTTTACTGATAGCCCTTCTTTTATCA 834

Qy        871 GCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAA 930
              |||||| | ||||  || || ||||| || || |||||  | || ||||||||||||||||
Db        835 GCTGAATCGGCCCAGAGGCCATCATATCCTTCCTCAAAATCCAAGAAAGACTGGGATAAA 894

Qy        931 CTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTG 990
              |||||||||||||| |||||||||||||||| || || ||||| || ||||||||||||||
Db        895 CTGGAAGCTGAAGTTAAAAAGGAGGAGAAGGAGGAGAAGCTTGAAGGCGATGCTGCATTG 954

Qy        991 AACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATG 1050
              |||||||| ||||||||||||||||  |||||||||||||| |||||  | || ||||||
Db        955 AACAAATTTTTCCGTGACATCTACAGTGATGCTGATGAAGACATGCGACGAGCAATGATG 1014

Qy       1051 AAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCA 1110
              || || || || ||||| || || |||||||| || ||||||||||||||||||  |
Db       1015 AAATCTTTTGTTGAATCTAACGGTACTGTTCTGTCGACCAATTGGAAAGATGTTGGCTCG 1074

Qy       1111 AAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              |||||||||| || ||||| |||||||| ||||| ||||| ||||||
Db       1075 AAGAAGGTAGAGGGAAGCCCACCTGATGGGATGGAGCTTAAGAAATGGGAGTACTAA 1131


RESULT 5
CF041238/c
LOCUS       CF041238                686 bp    mRNA    linear   EST 17-JUL-2003
DEFINITION  QCI23f02.yg QCI Zea mays cDNA clone QCI23f02, mRNA sequence.
ACCESSION   CF041238
VERSION     CF041238.1  GI:32936419
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="F2"
                     /db_xref="taxon:4577"
                     /clone="QCI23f02"
                     /tissue_type="embryo"
                     /clone_lib="QCI"
```

ORIGIN

Query Match 45.4%; Score 660.6; DB 6; Length 686;
Best Local Similarity 99.3%; Pred. No. 2.7e-160;
Matches 674; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
Qy        749 GCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGA 808
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        679 GCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGA 620

Qy        809 CATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGC 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 CATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGC 560

Qy        869 CAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATA 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATA 500

Qy        929 AACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCAT 988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 AACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCAT 440

Qy        989 TGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGA 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGA 380

Qy       1049 TGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAG 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAG 320

Qy       1109 CAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAA 1168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAA 260

Qy       1169 GTTTGGACTGCCCGTCTTTTGTAAATCCA-GGTCTTGGAAACTATGACCTAATTCTGCCC 1227
              |||||||||||||||||||||||||||| |||||||||||| |||||||||||||| ||
Db        259 GTTTGGACTGCCCGTCTTTTGTAAATCCAGGGTCTTGGAAACTGTGACCTAATTCTGTCC 200

Qy       1228 ACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGAC 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGAC 140

Qy       1288 AGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATT 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 AGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATT 80

Qy       1348 ATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAA 1407
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 ATTTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAA 20

Qy       1408 TGCCTATTCTTTTTGGTCT 1426
              ||||||||||||||||||||
Db          19 TGCCTATTCTTTTTGGTCT 1
```

RESULT 6
CA258233
LOCUS CA258233 772 bp mRNA linear EST 26-SEP-2003
DEFINITION SCCCRT3004D05.g RT3 Saccharum officinarum cDNA clone SCCCRT3004D05 5', mRNA sequence.
ACCESSION CA258233
VERSION CA258233.1 GI:35945956
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 772)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
Plate: 004 row: D column: 05
Seq primer: T7 Promoter Primer.
FEATURES Location/Qualifiers
source 1. .772
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCRT3004D05"
/lab_host="DH10B"
/clone_lib="RT3"
/note="Organ: Root apex from adult plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Root apex from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Query Match 45.2%; Score 657.2; DB 6; Length 772;
Best Local Similarity 93.7%; Pred. No. 2.1e-159;
Matches 697; Conservative 0; Mismatches 43; Indels 4; Gaps 1;

```
Qy         18 CACGAGAGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACGCCCAAAGCCGGAAGGGG 77
              | | |||||||||||||||||||||||||| || | | |     | || | |
Db         33 CGCCCGAGTCATCGTCTGGTCGTCGCCCGCAGACGCCCCAAGCC----GGAGGGGGTGC 88

Qy         78 CGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTT 137
              |    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CATGGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTT 148

Qy        138 CGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCT 197
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db        149 CGAGCTGGCCGCCGAGCTCTACACCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCT 208

Qy        198 CTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGA 257
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        209 CTACGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTTGCGGA 268

Qy        258 TGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGC 317
              |||||||||||||||| |||||||||| |||||||||||| ||||||||||||||||||
Db        269 TGCTAACAAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGC 328

Qy        318 TGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTC 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTC 388

Qy        378 TTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGC 437
              |||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db        389 CTATGCATCAGGCGATTCAAGGTTTACTCGCCTATTGAAGGAATGTGATGAGCGCATCGC 448

Qy        438 TGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGA 497
              ||||||||||||||||||||||||||||| ||||||||| |||||||||||||| ||||
Db        449 AGAGGAATCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGA 508

Qy        498 GGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAA 557
              |||||||||||||  |||||||||||||||||||| |||||| ||||||||||||||||
Db        509 GGACAAGGAGGATGCCACAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAA 568

Qy        558 ACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGC 617
```

```
Db        569 ACCTAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGC 628

Qy        618 TAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATC 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        629 TAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTAAGTGTATC 688

Qy        678 CATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTAT 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 CATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTAT 748

Qy        738 CCCTGAGAAATGCAAATATCAAGT 761
              |||||||||||||||||||||||||
Db        749 CCCTGAGAAATGCAAATATCAAGT 772


RESULT 7
AW076274
LOCUS       AW076274                668 bp    mRNA    linear   EST 14-OCT-1999
DEFINITION  614063D11.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW076274
VERSION     AW076274.1  GI:6031267
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614063  row: D  column: 11.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN

  Query Match             44.8%;  Score 652.4;  DB 2;  Length 668;
  Best Local Similarity   99.8%;  Pred. No. 3.6e-158;
  Matches  653;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        478 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 73

Qy        538 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 133

Qy        598 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 193

Qy        658 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 253

Qy        718 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTC 313

Qy        778 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 373

Qy        838 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 433

Qy        898 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 493

Qy        958 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 553

Qy       1018 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 613

Qy       1078 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 667


RESULT 8
CA176210/c
LOCUS       CA176210                798 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCJLST1027D01.b ST1 Saccharum officinarum cDNA clone SCJLST1027D01
            3', mRNA sequence.
ACCESSION   CA176210
VERSION     CA176210.1  GI:35105709
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 798)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 027  row: D  column: 01
            Seq primer: SP6 Promoter primer.
FEATURES             Location/Qualifiers
     source          1. .798
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCJLST1027D01"
                     /lab_host="DH10B"
                     /clone_lib="ST1"
                     /note="Organ: First apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
```

unidirectional cDNA library generated from [First apical stalk internodes of adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 44.7%; Score 650.8; DB 6; Length 798;
Best Local Similarity 91.5%; Pred. No. 1e-157;
Matches 732; Conservative 0; Mismatches 61; Indels 7; Gaps 4;

```
Qy        630 TGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCC 689
              ||||||||||||||||||||||| |||||| |||||||  |||||||||||||||| |
Db        798 TGCTGATAGTGTAGTCATTGATTNTGGTGACCAGATGTAAAGTGTATCCATTGAAGTCCN 739

Qy        690 TGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATG 749
              ||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db        738 TGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATATCCCCTGAGAAATG 679

Qy        750 CAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGAC 809
              || |||||||||| |||||||||||| || ||||||||||||||||||||||||||||||
Db        678 CANATATCAAGTCTCATCCACCAAGGTCGANATACGCCTTGCAAAAGCTGAGCAGGTGAC 619

Qy        810 ATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCC 869
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        618 ATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCC 559

Qy        870 AGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAA 929
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        558 AGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAA 499

Qy        930 ACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATT 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATT 439

Qy        990 GAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGAT 1049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGAT 379

Qy       1050 GAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGC 1109
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||| |
Db        378 GAAGTCATTCGTGGAATCTAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGATC 319

Qy       1110 AAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAG 1169
              ||||  ||| |||||||||| |||||||||||||||||| ||||||||||||||||||||
Db        318 AAAGGTGGTGGAAGGGAGCCCTCCTGATGGTATGGAGCTTAAGAAGTGGGAATACTAAAG 259

Qy       1170 TTTGGAC-TGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCA 1228
              || |||| |||||    |||||||||||||||||||| || ||||||||| |   | |||
Db        258 TTCGGACTTGCCCCGTCTTTGTAAATCCAGGTCTTGGGAATTATGACCTAGTAACGTCCA 199

Qy       1229 CCCA--TAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGA 1286
              ||||  | ||||||||||||||   | | |    |||||||||||||||||||||||||
Db        198 CCCATTTGGTGCCATGAGCTTG---GTTAAGTCTTCTGCTTTTGTAAGCTTTCTGTATAT 142

Qy       1287 CAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACAT-AGGTATGCTA 1345
              |   |   | |||||||||||||||||||||||||| ||||||||||| |||||||||
Db        141 GACAGTCTAGCGTGTCTGAGTGCTACTTGCCGTAGTTTTTGATGGACATCTGGTATGCTA 82

Qy       1346 TTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTATGG 1405
              || |||||||||||  |||| |||||||||| |||||||||||||||||||| | |||
Db          81 TTCTGTGGTATCGACTGCGAAACATCTGTGGTAACAAAGACAGCAAACTATAATCTGTGG 22

Qy       1406 AATGCCTATTCTTTTTGGTC 1425
              ||||| ||||||||||||||
Db          21 AATGCTTATTCTTTTTGGTC 2
```

RESULT 9

```
CA184150
LOCUS       CA184150                 756 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCRFST3142H05.g ST3 Saccharum officinarum cDNA clone SCRFST3142H05
            5', mRNA sequence.
ACCESSION   CA184150
VERSION     CA184150.1  GI:35121633
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 756)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 142  row: H  column: 05
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .756
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCRFST3142H05"
                     /lab_host="DH10B"
                     /clone_lib="ST3"
                     /note="Organ: Fourth apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Fourth apical
                     stalk internodes of adult plants]. cDNA was prepared from
                     polyA+ mRNA using SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN
```

Query Match 44.5%; Score 647; DB 6; Length 756;
Best Local Similarity 95.0%; Pred. No. 9.6e-157;
Matches 721; Conservative 0; Mismatches 32; Indels 6; Gaps 5;

```
Qy         85 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 62

Qy        145 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 204
              | | |||||  ||||| |||||||||||||||||||||||||||||||||||||| |||
Db         63 GGCGCCGAG--CTACACCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTACGCC 120

Qy        205 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 264
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        121 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTTGCGGATGCTAAC 180

Qy        265 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 324
              |||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||||
Db        181 AAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGC 240

Qy        325 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 384
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCA 300
```

```
Qy        385 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 444
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        301 TCAGGCGATTCAAGGTTTACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 360

Qy        445 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAG 504
              |||||||||||||||||||||||| ||||||||| ||||||||||||| ||||||||||
Db        361 TCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAG 420

Qy        505 GAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAA 564
              ||||||  |||||||||||||||||||| |||||| |||||||||||||||||| |||
Db        421 GAGGATGCCACAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAA 480

Qy        565 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGT 624
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        481 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGT 540

Qy        625 GTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAA 684
              |||||||||||||||||||||||||  |||||||||||||||| |||||||||||||| |
Db        541 GTTCCTGCTGATAGTGTAGTCATTGGATTTGGTGAACAGATGTTAAGTGTATCCATTGGA 600

Qy        685 GTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAG 744
              || |||||||||||||||||||| |||||||||||| ||||||||| ||||||||||
Db        601 GTCCCTGGTGAAGAACCATACCATTNTCAGCCCCGTCTGGTTTCTAAGAATATCCCTGAG 660

Qy        745 AAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAG 804
              ||||| | |||||||||||||||||| | |||| |||||||||||||| |||||||||| |
Db        661 AAATG-ANATATCAAGTCTTATCCAACCAGGTCGAAATACGCCTTGC-AAAGCTGAGC-G 717

Qy        805 GTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 843
              |||||||||||| ||||||| ||||||||||||||||||
Db        718 GTGACATGGACACCCCTGGA-TATAGTGGAAGACCAAAG 755
```

RESULT 10
AW090936

```
LOCUS       AW090936                659 bp    mRNA    linear   EST 18-OCT-1999
DEFINITION  614068G07.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW090936
VERSION     AW090936.1  GI:6056546
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614068  row: G  column: 07.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN
```

```
  Query Match           44.2%;  Score 643.4;  DB 2;  Length 659;
  Best Local Similarity  99.8%;  Pred. No. 7.9e-156;
  Matches  644;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        321 ATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTA 74

Qy        381 TGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGA 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGA 134

Qy        441 GGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGA 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGA 194

Qy        501 CAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACC 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACC 254

Qy        561 AAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAA 314

Qy        621 GGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCAT 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCAT 374

Qy        681 TGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCC 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCC 434

Qy        741 TGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGA 800
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        435 TGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGA 494

Qy        801 GCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGAT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 GCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGAT 554

Qy        861 AAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGA 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGA 614

Qy        921 CTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGA 965
              |||||||||||||||||||||||||||||||||||||||||||||
Db        615 CTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGA 659
```

RESULT 11
CA245820

```
LOCUS       CA245820                759 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCEZFL5084D06.g Saccharum officinarum FL5 Saccharum officinarum
            cDNA clone SCEZFL5084D06 5', mRNA sequence.
ACCESSION   CA245820
VERSION     CA245820.1  GI:35324833
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
```

```
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 084  row: D  column: 06
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCEZFL5084D06"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL5"
                     /note="Organ: Developed inflorescence (20cm-long) without
                     rachis; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Developed
                     inflorescence  (20cm-long) without rachis]. cDNA was
                     prepared from polyA+ mRNA using  SuperScript Plasmid
                     System Kit (Invitrogen). The double-strand cDNAs were
                     fractionated in a sepharose CL-2B 40cm-columns and
                     fragments sizing  between 0.8 and 1.5 Kb were
                     directionally cloned into the vector. Details  of each
                     source of RNA and library construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             44.1%;  Score 641;  DB 6;  Length 759;
  Best Local Similarity   93.6%;  Pred. No. 3.5e-155;
  Matches  700;  Conservative    0;  Mismatches   43;  Indels    5;  Gaps    3;

Qy        288 GATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 347
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 61

Qy        348 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCG 407
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||||
Db         62 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTTACTCG 121

Qy        408 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 181

Qy        468 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 527
              |||||||| |||||||||||||| |||||||||||||||  ||||||||||||||||||
Db        182 GGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATGGAGAA 241

Qy        528 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 587
              |||||| |||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db        242 TACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTACAACAG 301

Qy        588 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 647
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 361

Qy        648 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 707
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db        362 TGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCA 421

Qy        708 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 767
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        422 TTTTCAGCCCCGTCTGTTTTCTAAGAATATCCCTGAGAAATGCAAATATCAAGTCTTATC 481

Qy        768 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 827
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 541

Qy        828 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 887
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        542 TAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 601

Qy        888 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGG---GATAAACTGGAAGCTGAAGT 944
              |||||||||||||||||||||||| ||||||     |   |||||   |||||||||
Db        602 ACCTTCATACCCATCTTCAAAGGCAAAAAAAAGACTGGNNGATAACTGGGAAGCTGAGTC 661

Qy        945 CAAAAAGGAGGAGAA-GGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCC 1003
              |||| ||| ||||| ||||||| |||||||||||||||||  ||||    ||||| |||
Db        662 AAAAAGGGAAGAGAAGGGAAGAANAACTTGATGGTGATGCTTGATTGNNACAATTC-TCC 720

Qy       1004 GTGACATCTACAAGGATGCTGATGAAGA 1031
              |||||||||| | ||||||||||||
Db        721 GTGACATCTAACAAGATGCTGATGAAAA 748


RESULT 12
CA227695
LOCUS       CA227695                778 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCJLFL3014D05.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCJLFL3014D05 5', mRNA sequence.
ACCESSION   CA227695
VERSION     CA227695.1  GI:35288469
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 778)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 014  row: D  column: 05
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .778
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCJLFL3014D05"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             43.8%;  Score 637.4;  DB 6;  Length 778;
  Best Local Similarity   91.9%;  Pred. No. 3e-154;
  Matches  718;  Conservative    0;  Mismatches   56;  Indels    7;  Gaps    4;

Qy        650 ATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATT 709
              |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db          1 ATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCATT 60

Qy        710 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 769
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCA 120

Qy        770 CCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 829
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATA 180

Qy        830 GTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 889
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGAC 240

Qy        890 CTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 949
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        241 CTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAA 300

Qy        950 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACA 360

Qy       1010 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAA 1069
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        361 TCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCTA 420

Qy       1070 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCC 1129
              |||||||||||||||||||||||||||||||||||||| |||||  ||| ||||||||||
Db        421 ATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGGTGGTGGAAGGGAGCC 480

Qy       1130 CCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGAC-TGCCCGTCTTTT 1188
              | ||||||||||||||||| |||||||||||||||||||||| |||| |||||    |||
Db        481 CTCCTGATGGTATGGAGCTTAAGAAGTGGGAATACTAAAGTTCGGACTTGCCCCGTCTTT 540

Qy       1189 GTAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCA--TAGTGCCATGAGCT 1246
              ||||||||||||||||| || ||||||||| |   | |||||||  | ||||||||||||
Db        541 GTAAATCCAGGTCTTGGGAATTATGACCTAGTAACGTCCACCCATTTGGTGCCATGAGCT 600

Qy       1247 TGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAG 1306
              ||   | | |    ||||||||||||||||||||||||   |   |   | |||||||||
Db        601 TG---GTTAAGTCTTCTGCTTTTGTAAGCTTTCTGTATATGACAGTCTAGCGTGTCTGAG 657

Qy       1307 TGCTACTTGCCGTAGCTTTTGATGGACAT-AGGTATGCTATTATGTGGTATCGACCCCGA 1365
              ||||||||||||||  |||||||||||||  ||||||||||| ||||||||||||  |||
Db        658 TGCTACTTGCCGTAATTTTTGATGGACATCTGGTATGCTATTCTGTGGTATCGACTGCGA 717

Qy       1366 ACCATCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTC 1425
              | ||| |  |||| | |||||||||||||||||| | ||||||| |||||||||||||||
Db        718 AACATTTTGGGTAACCAAGACAGCAAACTATAATCTGTGGAATGGCTATTCTTTTTGGTC 777

Qy       1426 T 1426
              |
Db        778 T 778
```

```
RESULT 13
CA215140
LOCUS       CA215140                 666 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCSBAD1128H12.g AD1 Saccharum officinarum cDNA clone SCSBAD1128H12
            5', mRNA sequence.
ACCESSION   CA215140
VERSION     CA215140.1  GI:35263122
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol..24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 128  row: H  column: 12
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1..666
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCSBAD1128H12"
                     /lab_host="DH10B"
                     /clone_lib="AD1"
                     /note="Organ: seedlings inoculated with Gluconacetobacter
                     diazotroficans; Vector: pSport1; Site_1: SalI; Site_2:
                     NotI; An unidirectional cDNA library generated from
                     [seedlings inoculated with  Gluconacetobacter
                     diazotroficans]. cDNA was prepared from polyA+ mRNA  using
                     SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             43.4%;  Score 631;  DB 6;  Length 666;
  Best Local Similarity   96.5%;  Pred. No. 1.3e-152;
  Matches  643;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy        420 ATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTAC 479
              |||||||||| ||||||||||||||||||||||||||||||||||||| ||||||||| |
Db          1 ATGTGATGAGTGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCC 60

Qy        480 TGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGT 539
              ||||||||||||| |||||||||||||||||  ||||||||||||||||||||| |||||
Db         61 TGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATGGAGAATACACCGCCAGT 120

Qy        540 GATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGT 599
              | ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 GGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTACAACAGTGCCACAGAAGT 180

Qy        600 GGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGA 659
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGA 240

Qy        660 ACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCG 719
              ||||||||| ||||||||||||||||| ||||||||||||||||||||||||||||| ||
Db        241 ACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCATTTTCAGCCACG 300

Qy        720 TCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGA 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        301 TCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGA 360

Qy        780 AATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACC 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACC 420

Qy        840 AAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCC 899
              |||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGGCTGTTCCCCAAAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCC 480

Qy        900 ATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAA 959
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAA 540

Qy        960 GGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGA 1019
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        541 GGAAGAANAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACANGGA 600
```

```
Qy       1020 TGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGT 1079
              ||||||||||||||||||  |||||||||||||||||||||||| ||| |||||||
Db        601 TGCTGATGAAGATATGCGGNAGGCCATGATGAAGTCATTCGTGGAATCTAATTGCACTGT 660

Qy       1080 TCTCTC 1085
              ||||||
Db        661 TCTCTC 666


RESULT 14
AW076372
LOCUS       AW076372                681 bp    mRNA    linear   EST 14-OCT-1999
DEFINITION  614066C10.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW076372
VERSION     AW076372.1  GI:6031470
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614066  row: C  column: 10.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN

  Query Match             43.2%;  Score 628.8;  DB 2;  Length 681;
  Best Local Similarity   99.7%;  Pred. No. 5e-152;
  Matches  630;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        735 TATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAA 794
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         19 TATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAA 78

Qy        795 AGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 AGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCA 138

Qy        855 GAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAA 914
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        139 GAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCATAGGCGAA 198

Qy        915 AAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGA 258

Qy        975 TGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        259 TGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATAT 318

Qy       1035 GCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTG 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTG 378

Qy       1095 GAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAA 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 GAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAA 438

Qy       1155 GTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGA 498

Qy       1215 CCTAATTCTGCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CCTAATTCTGCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAG 558

Qy       1275 CTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACA 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACA 618

Qy       1335 TAGGTATGCTATTATGTGGTATCGACCCCGAA 1366
              ||||||||||||||||||||||||||||||||
Db        619 TAGGTATGCTATTATGTGGTATCGACCCCGAA 650


RESULT 15
AW155786
LOCUS       AW155786                658 bp    mRNA    linear   EST 04-NOV-1999
DEFINITION  614096F11.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW155786
VERSION     AW155786.1  GI:6227082
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614096  row: F  column: 11.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN

  Query Match             43.1%;  Score 627;  DB 2;  Length 658;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-151;
  Matches  627;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        829 AGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGA 888
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 AGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGA 68

Qy        889 CCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAA 948
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAA 128

Qy        949 AAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGAC 1008
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 AAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGAC 188

Qy       1009 ATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCA 1068
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 ATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCA 248

Qy       1069 AATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGC 1128
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 AATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGC 308

Qy       1129 CCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTT 1188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 CCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTT 368

Qy       1189 GTAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTG 1248
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GTAAATCCAGGTCTTGGAAACTATGACCTAATTCTGCCCACCCATAGTGCCATGAGCTTG 428

Qy       1249 CTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTG 1308
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGTATGACAGTCTAGCGTGTGTCTGAGTG 488

Qy       1309 CTACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACC 1368
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 CTACTTGCCGTAGCTTTTGATGGACATAGGTATGCTATTATGTGGTATCGACCCCGAACC 548

Qy       1369 ATCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCTAA 1428
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 ATCTGTGGTATCAAAGACAGCAAACTATAATGTATGGAATGCCTATTCTTTTTGGTCTAA 608

Qy       1429 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1455
                ||||||||||||||||||||||||||||
Db        609 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 635
```

Search completed: April 23, 2005, 05:43:21
Job time : 5238.64 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 22, 2005, 23:55:26 ; Search time 841.735 Seconds
(without alignments)
10232.696 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcggcacgagcggcac.....aaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	734.4	50.5	1104	12	ADQ37150	Adq37150 Cell prol
2	734.4	50.5	1104	12	ADQ15734	Adq15734 Rice stre
3	571	39.2	571	10	ADF51379	Adf51379 Maize Rar
4	420	28.9	1366	3	AAC39274	Aac39274 Arabidops
5	419.6	28.8	1324	3	AAC40498	Aac40498 Arabidops
6	418.4	28.8	1361	3	AAC51074	Aac51074 Arabidops
7	417.6	28.7	1077	6	ADG87792	Adg87792 A. thalia
8	393.2	27.0	1117	3	AAC42413	Aac42413 Arabidops
9	388	26.7	388	10	ADF51380	Adf51380 Maize Rar
10	286	19.7	287	6	ABL71197	Abl71197 Corn tass
c 11	240.2	16.5	581	9	ACL25190	Acl25190 DNA clone
12	221.8	15.2	585	13	ACN47711	Acn47711 Cotton pr
13	217.8	15.0	534	13	ACN59636	Acn59636 Cotton gy
14	216.4	14.9	272	6	ABL75330	Abl75330 Corn tass
c 15	214.4	14.7	541	9	ACL25189	Acl25189 DNA clone
c 16	203.6	14.0	571	9	ACL25186	Acl25186 DNA clone
17	201.6	13.9	291	6	ABL73761	Abl73761 Corn tass
18	194.4	13.4	528	13	ACN59570	Acn59570 Cotton gy
c 19	189.6	13.0	500	9	ACL25185	Acl25185 DNA clone
c 20	183.8	12.6	573	13	ACN47765	Acn47765 Cotton pr
c 21	175.6	12.1	523	13	ACN61180	Acn61180 Cotton gy
22	164.4	11.3	272	10	ABX86113	Abx86113 Corn ear-
23	163	11.2	1535	8	ACD13363	Acd13363 Human DNA
24	163	11.2	1555	10	ADD29650	Add29650 Human tum
25	163	11.2	1776	5	ABV28161	Abv28161 Human pro
26	163	11.2	1776	5	ABV22337	Abv22337 Human pro
27	156.4	10.7	1025	3	AAZ52451	Aaz52451 HTRM clon
28	156.4	10.7	1185	6	ABL42163	Abl42163 Nucleotid
29	156.4	10.7	1264	6	ABQ55038	Abq55038 Human ova
30	156.4	10.7	1366	8	ABZ24700	Abz24700 Human cel
31	156.4	10.7	2053	2	AAX27345	Aax27345 Human sec
32	156.4	10.7	2053	9	ADA07224	Ada07224 Human cDN
33	156.4	10.7	2053	12	ADN40960	Adn40960 Novel hum
c 34	142.2	9.8	409	8	ABX62506	Abx62506 Arabidops
35	133.6	9.2	615	13	ADQ55550	Adq55550 Novel can
36	129.2	8.9	481	10	ADE82158	Ade82158 Arabidops
c 37	111.4	7.7	464	10	ADB50274	Adb50274 Primary r
c 38	111.2	7.6	340	9	ACL25187	Acl25187 DNA clone
c 39	111.2	7.6	412	9	ACL25184	Acl25184 DNA clone
40	93.2	6.4	672	12	ADJ40154	Adj40154 Plant cDN
41	90.8	6.2	1242	6	ABZ31779	Abz31779 Candida a
42	85.4	5.9	431	6	ABL62696	Abl62696 Colon ade
43	78.2	5.4	518	6	ABQ55648	Abq55648 Human ova
44	78	5.4	1303	12	ADJ43595	Adj43595 Plant cDN
45	75.8	5.2	869	4	AAK63551	Aak63551 Human imm

ALIGNMENTS

RESULT 1
ADQ37150
ID ADQ37150 standard; DNA; 1104 BP.
XX
AC ADQ37150;
XX
DT 07-OCT-2004 (first entry)
XX
DE Cell proliferation-related nucleic acid sequence #155.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response; ds.
XX
OS Oryza sativa.
XX
PN WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
PS Claim 57; SEQ ID NO 309; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body

```
CC   of the specification as EPO data.
XX
SQ   Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

  Query Match           50.5%;  Score 734.4;  DB 12;  Length 1104;
  Best Local Similarity  80.5%;  Pred. No. 3.3e-151;
  Matches  893;  Conservative    0;  Mismatches  196;  Indels   21;  Gaps    2;

Qy         73 AGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGAC 132
              | ||    ||  ||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 ATGGCAACCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGAC 60

Qy        133 GACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCC 192
              |||||||||||| ||| ||||||||||||  ||||| ||||| ||| | ||||||||||||
Db         61 GACTTCGAGCTCGCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCC 120

Qy        193 GACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTG 252
              || ||||| |||||||||||||||||| || |||||||| ||||||||||||||||||||
Db        121 GAGCTCTACGCCGACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTA 180

Qy        253 GCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAA 312
              || |||||||||| || ||||| ||||| ||     ||||| || |||||    || |||
Db        181 GCTGATGCTAACAAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAA 240

Qy        313 GGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGT 372
              || |||||||| ||     || || || || ||||||||||| || |||||||| ||||||
Db        241 GGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTGGGT 300

Qy        373 TCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGC 432
              |  || |  ||||| || || ||||||||| |||| ||| ||||||| ||||||||||||
Db        301 TACTCGTTCGCATCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGC 360

Qy        433 ATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGG----------------CT 477
              || ||||||||   ||   | |  || || ||||| | ||| |                |
Db        361 ATTGCTGAGGAGCTTACTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCCCCC 420

Qy        478 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 537
              |||| ||| ||  | ||||||| ||||| |||    |||| ||||| ||||||||||||
Db        421 TCTGTTGCTTCTTTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCA 480

Qy        538 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 597
              || ||||      |   || |||||||||||||| |||| |||||||||||| ||||||
Db        481 ATGGTAGA------AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAA 534

Qy        598 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 657
              || ||| ||||||| ||||| |||||||||||||||| | ||| ||  ||||||||||||
Db        535 GTTGTATTGACAATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGT 594

Qy        658 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 717
              ||||| ||||| ||||| || |||||||| ||||| || || || ||||||||||||||
Db        595 GAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCT 654

Qy        718 CGTCTGTTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTT 777
              |||||||||||||||| |||||||||||| ||| ||| |||||  ||||||| ||||||
Db        655 CGTCTGTTTTTCTAAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTT 714

Qy        778 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 837
              |||||| | || || |||||||| ||| | |||||||  | || ||||||  |  || |
Db        715 GAAATAAGACTGGCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAA 774

Qy        838 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 897
              |||||| ||||||| || ||||||| | | ||||||||| | ||||  || || |||||
Db        775 CCAAAGGCTGTTCCACAAAAGATAATCCCTCCAGCTGAATCGGCCCAGAGGCCATCATAT 834

Qy        898 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 957
              || || |||||  | || ||||||||||||||||||||||||||||| ||||||||||||
Db        835 CCTTCCTCAAAATCCAAGAAAGACTGGGATAAACTGGAAGCTGAAGTTAAAAAGGAGGAG 894

Qy        958 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 1017
              ||||| || || ||||| || |||||||||||||||||||||| ||||||||||||||||
Db        895 AAGGAGGAGAAGCTTGAAGGCGATGCTGCATTGAACAAATTTTTCCGTGACATCTACAGT 954

Qy       1018 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 1077
              |||||||||||||| |||||  | || |||||||| || || || ||||| || || |||
Db        955 GATGCTGATGAAGACATGCGACGAGCAATGATGAAATCTTTTGTTGAATCTAACGGTACT 1014

Qy       1078 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGAT 1137
              ||||| || |||||||||||||||||||  | |||||||||| || ||||| ||||||
Db       1015 GTTCTGTCGACCAATTGGAAAGATGTTGGCTCGAAGAAGGTAGAGGGAAGCCCACCTGAT 1074

Qy       1138 GGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              || |||||||| ||||| ||||| ||||||
Db       1075 GGGATGGAGCTTAAGAAATGGGAGTACTAA 1104


RESULT 2
ADQ15734
ID   ADQ15734 standard; DNA; 1104 BP.
XX
AC   ADQ15734;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Rice stress-related protein coding sequence #72.
XX
KW   rice; stress-related protein; plant maturation; plant development;
KW   plant proliferation; plant senescence; plant disease-resistance;
KW   plant stress response; transgenic plant; pest tolerance;
KW   herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
KW   improved nutritional value; increased yield; increased proliferation;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO2004061080-A2.
XX
PD   22-JUL-2004.
XX
PF   23-DEC-2003; 2003WO-US041098.
XX
PR   26-DEC-2002; 2002US-0436564P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Cooper B;
XX
DR   WPI; 2004-534374/51.
DR   P-PSDB; ADQ15735.
XX
PT   New isolated nucleic acids and proteins, useful for producing transgenic
PT   plants having improved properties, e.g. tolerance to pests, herbicides,
PT   or biotic or abiotic stresses, improved nutritional value, or increased
PT   yield or proliferation.
XX
PS   Claim 45; SEQ ID NO 143; 551pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of rice
CC   stress-related proteins. The DNA and protein sequences of the invention
CC   are useful for regulating and controlling plant maturation and
CC   development, including proliferation, senescence, disease-resistance, or
CC   stress response. They are also useful for producing transgenic plants
CC   having improved properties, e.g. tolerance to pests, herbicides, or
CC   biotic or abiotic stresses, improved nutritional value, increased yield
CC   or proliferation, or improved structure causing less loss from lodging or
CC   shattering. The present DNA sequence encodes a rice stress-related
CC   protein of the invention.
XX
SQ   Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

  Query Match           50.5%;  Score 734.4;  DB 12;  Length 1104;
  Best Local Similarity  80.5%;  Pred. No. 3.3e-151;
  Matches  893;  Conservative    0;  Mismatches  196;  Indels   21;  Gaps    2;

Qy         73 AGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGAC 132
              | ||    ||  ||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 ATGGCAACCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGAC 60
```

```
Qy        133 GACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCC 192
              |||||||||| ||| |||||||||||  ||||| ||||| ||| | ||||||||||||
Db         61 GACTTCGAGCTCGCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCC 120

Qy        193 GACCTCTATGCCGACCGCGCCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTG 252
              || ||||| |||||||||||||||| || |||||||| |||||||||||||||||||||
Db        121 GAGCTCTACGCCGACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTA 180

Qy        253 GCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAA 312
              || |||||||||| || ||||| ||||| ||     ||||| || |||||     || |||
Db        181 GCTGATGCTAACAAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAA 240

Qy        313 GGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGT 372
              || |||||||| ||     || || || || |||||||||| || |||||||| ||||||
Db        241 GGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTGGGT 300

Qy        373 TCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGC 432
              |  || |  ||||| || || |||||||||| |||| ||| ||||||| ||||||||||||
Db        301 TACTCGTTCGCATCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGC 360

Qy        433 ATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGG----------------CT 477
              || |||||||   ||    | |  || || ||||| | ||| |                |
Db        361 ATTGCTGAGGAGCTTACTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCCCCC 420

Qy        478 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 537
              |||| ||| ||  | ||||||| ||||| |||    |||| ||||| ||||||||||||
Db        421 TCTGTTGCTTCTTTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCA 480

Qy        538 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 597
              || ||||      |   || ||||||||||||| |||| |||||||||| ||||||
Db        481 ATGGTAGA------AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAA 534

Qy        598 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 657
              || ||| ||||||| ||||| |||||||||||||||| | ||| ||  ||||||||||
Db        535 GTTGTATTGACAATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGT 594

Qy        658 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 717
              ||||| ||||| ||||| || |||||||| ||||| || || || ||||||||||||||
Db        595 GAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCT 654

Qy        718 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTT 777
              |||||||||||||||| |||||||||||| ||| ||| |||||  ||||||| ||||||
Db        655 CGTCTGTTTTCTAAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTT 714

Qy        778 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 837
              |||||| | || || |||||||| ||| | |||||||  | || ||||||  |  || |
Db        715 GAAATAAGACTGGCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAA 774

Qy        838 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 897
              |||||| ||||||| || |||||||  | | ||||||||| | ||||  || || |||||
Db        775 CCAAAGGCTGTTCCACAAAAGATAATCCCTCCAGCTGAATCGGCCCAGAGGCCATCATAT 834

Qy        898 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 957
              || || |||||  | || |||||||||||||||||||||||||||| ||||||||||||
Db        835 CCTTCCTCAAAATCCAAGAAAGACTGGGATAAACTGGAAGCTGAAGTTAAAAAGGAGGAG 894

Qy        958 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 1017
              ||||| || || ||||| || |||||||||||||||||| ||||||||||||||||
Db        895 AAGGAGGAGAAGCTTGAAGGCGATGCTGCATTGAACAAATTTTTCCGTGACATCTACAGT 954

Qy       1018 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 1077
              |||||||||||||| |||||  | || |||||||| || || || ||||| || || |||
Db        955 GATGCTGATGAAGACATGCGACGAGCAATGATGAAATCTTTTGTTGAATCTAACGGTACT 1014

Qy       1078 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGAT 1137
              ||||| || |||||||||||||||  |||||||||| || ||||| ||||||
Db       1015 GTTCTGTCGACCAATTGGAAAGATGTTGGCTCGAAGAAGGTAGAGGGAAGCCCACCTGAT 1074

Qy       1138 GGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              || |||||||| ||||| ||||| ||||||
Db       1075 GGGATGGAGCTTAAGAAATGGGAGTACTAA 1104
```

RESULT 3
ADF51379
ID ADF51379 standard; cDNA; 571 BP.
XX
AC ADF51379;
XX
DT 12-FEB-2004 (first entry)
XX
DE Maize Rar interactior 2 (RarInt2) DNA seq id 2.
XX
KW plant protectant; gene therapy; Rar1; disease resistance; plant;
KW immunogen; maize; Rar interactor 2; RarInt2; gene; ss.
XX
OS Zea mays.
XX
PN US2003167504-A1.
XX
PD 04-SEP-2003.
XX
PF 12-FEB-2002; 2002US-00074473.
XX
PR 12-FEB-2001; 2001US-0268157P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Crane EH;
XX
DR WPI; 2003-898130/82.
XX
PT New maize Rar1-interactor nucleic acid useful for enhancing disease
PT resistance in plants, as a probe or amplification primer, in recombinant
PT expression of polypeptides, or as an immunogen in the preparation or
PT screening of antibodies.
XX
PS Claim 1; SEQ ID NO 2; 30pp; English.
XX
CC The invention describes an isolated nucleic acid capable of interacting
CC with Rar1. The nucleic acid comprises a member selected from: (a) a
CC polynucleotide comprising a sequence of 388 (S1), 571 (S2) or 691 (S3) bp
CC fully defined in the specification; (b) a polynucleotide amplified from a
CC Zea mays nucleic acid library using primers that selectively hybridise,
CC under high stringency conditions, to loci within the polynucleotide in
CC (a); (c) a polynucleotide encoding a maize Rar1-interactor protein; (d) a
CC polynucleotide having at least 80 or 90% sequence identity to (a); (e) a
CC polynucleotide which hybridises under high stringency conditions to (a);
CC and (f) a polynucleotide complementary to any of the above
CC polynucleotides. The nucleic acid molecule and methods are useful in
CC enhancing disease resistance in plants. The nucleic acid molecule may
CC also be used as a probe or an amplification primer, in recombinant
CC expression of polypeptides, or as an immunogen in the preparation and/or
CC screening of antibodies. This sequence represents a maize Rar interactor
CC 2 (RarInt2) polynucleotide.
XX
SQ Sequence 571 BP; 151 A; 154 C; 161 G; 105 T; 0 U; 0 Other;

Query Match 39.2%; Score 571; DB 10; Length 571;
Best Local Similarity 100.0%; Pred. No. 2e-115;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAATTCGGCACGAGCGGCACGAGAGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAATTCGGCACGAGCGGCACGAGAGTCATCGTCTGGTCGTCGCCCGCAGAGGCGCAGACG 60

Qy         61 CCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCC 120

Qy        121 TTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGG 180

Qy        181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240

Qy        241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300

Qy        301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360

Qy        361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420

Qy        421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480

Qy        481 GTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTG 540

Qy        541 ATAGAACCCCCAAGCAAACCAAAATACAGGC 571
                ||||||||||||||||||||||||||||||
Db        541 ATAGAACCCCCAAGCAAACCAAAATACAGGC 571


RESULT 4
AAC39274
ID   AAC39274 standard; DNA; 1366 BP.
XX
AC   AAC39274;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 24023.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
```

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 28.9%; Score 420; DB 3; Length 1366;
Best Local Similarity 63.0%; Pred. No. 3.7e-82;
Matches 683; Conservative 0; Mismatches 395; Indels 6; Gaps 2;

```
Qy        88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
             |||  |||  | |     || || || || || ||  | || || ||||||||  | ||
Db        60 GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 119

Qy       148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
               ||  | |||  | | ||||| |||   |  |||     |||||  | |||  |||||
Db       120 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 179

Qy       208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
             || || |||||  |||||||  |||  |||| |||||| |||||  |  ||||| ||||||
Db       180 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 239

Qy       268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
             || |||||| | || || | | ||    ||||| ||   | | || ||  |||| || ||
Db       240 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 299

Qy       328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
             ||||| ||||||||     ||||| || || || || ||    |||  |||||  |||| |
Db       300 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 359

Qy       388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
               ||  ||| ||||        |  |  | ||||| |||   || || || || |||
Db       360 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 419

Qy       448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
             |   | |     ||  ||     |    | | | || || | |   |
Db       420 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476

Qy       508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
             |       | |   |    | |    ||| |    | | | ||   ||||||  | | |
Db       477 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 536

Qy       568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
             ||||| || | |||| | |   |  |||||  ||| |||||||  || || ||   ||
Db       537 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 596

Qy       628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
             |||   || |  |||    | || |||||||| |||||  |||||||   |||||| |||
Db       597 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 656

Qy       688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
             |||| || ||| | || ||| | |||||  |  ||||     ||||| || || |||||
Db       657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 716

Qy       748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
             |||| || | |||| || || ||||| ||||| || || ||||| ||||| |||    |
Db       717 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 776

Qy       808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
             || ||| |  |||| || ||| ||  | | | ||    | | |||  |    |  | |
Db       777 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 836

Qy       868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
               ||||         | |  |||||    |||||||||||  ||| |   ||||||||||||
Db       837 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893

Qy       928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
```

```
                ||  |||||||||||||| || ||  |||||||||| || || |||||||| ||||| ||
Db        894 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 953

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
              |||||||||| ||| | ||||| |||  |  ||| |||||||| ||| || |||| |||
Db        954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
              |  || ||||| |  || || ||||| || || || || || || |||||||| |||||
Db       1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
               | ||||| || ||  | |  || || ||||| |||||||||||||||||||| || | |
Db       1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133

Qy       1168 AGTT 1171
                ||
Db       1134 TCTT 1137
```

RESULT 5
AAC40498
ID AAC40498 standard; DNA; 1324 BP.
XX
AC AAC40498;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28522.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 28.8%; Score 419.6; DB 3; Length 1324;
Best Local Similarity 62.9%; Pred. No. 4.4e-82;
Matches 682; Conservative 1; Mismatches 395; Indels 6; Gaps 2;

```
Qy        88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
             ||| ||| | |      || || || || || :| | || || ||||||||  | ||
Db        60 GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTYTTCTAGATGATGACTTCGATGTTGCT 119

Qy       148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
               || | |||  | | ||||| |||   |  |||      ||||| | |||  |||||
Db       120 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 179

Qy       208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
             || || |||||  |||||||  |||  |||| |||||| ||||| |  ||||| ||||||
Db       180 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 239

Qy       268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
             || |||||| | || || | | ||   ||||| ||   | | || ||  |||| || ||
Db       240 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 299

Qy       328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
             ||||| ||||||||    ||||| || || || || ||   |||  |||||  |||| |
Db       300 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 359

Qy       388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
               ||  ||| ||||       |  |  | ||||| |||   || || || || |||
Db       360 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 419

Qy       448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
             |   | |      ||  ||     | | | || || | |    |     |
Db       420 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476

Qy       508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
             |       |  |   |  | |    ||| |    | | | ||   |||||| |  | |
Db       477 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 536

Qy       568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
             ||||| || | |||| | |   |  ||||| |||  ||||||| || || ||    ||
Db       537 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 596

Qy       628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
             |||   || |  |||    | || |||||||| |||||  |||||||   |||||| |||
Db       597 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 656

Qy       688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
              |||| || ||| | || ||| | |||||  |  ||||     ||||| || || |||||
Db       657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 716

Qy       748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
             |||| || | |||| || || ||||| ||||| || || ||||| ||||| |||    |
Db       717 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 776

Qy       808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
             || ||| |  |||| || ||| || | | ||    | | ||| |     |  | |
Db       777 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 836

Qy       868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
              ||||         | |  |||||     |||||||||||| ||| |   ||||||||||||
Db       837 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893

Qy       928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
             ||  |||||||||||||| || ||  |||||||||| || || |||||||| ||||| ||
Db       894 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 953

Qy       988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
             |||||||||| ||| | ||||| |||  |  ||| |||||||| ||| || |||| |||
Db       954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy      1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
```

```
                |  || |||||  |  || || ||||| || || || || || || |||||||| |||||
Db       1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
                | ||||| || ||  | |  || || ||||| ||||||||||||||||||| || | |
Db       1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133

Qy       1168 AGTT 1171
                ||
Db       1134 TCTT 1137
```

RESULT 6
AAC51074
ID AAC51074 standard; DNA; 1361 BP.
XX
AC AAC51074;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67179.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

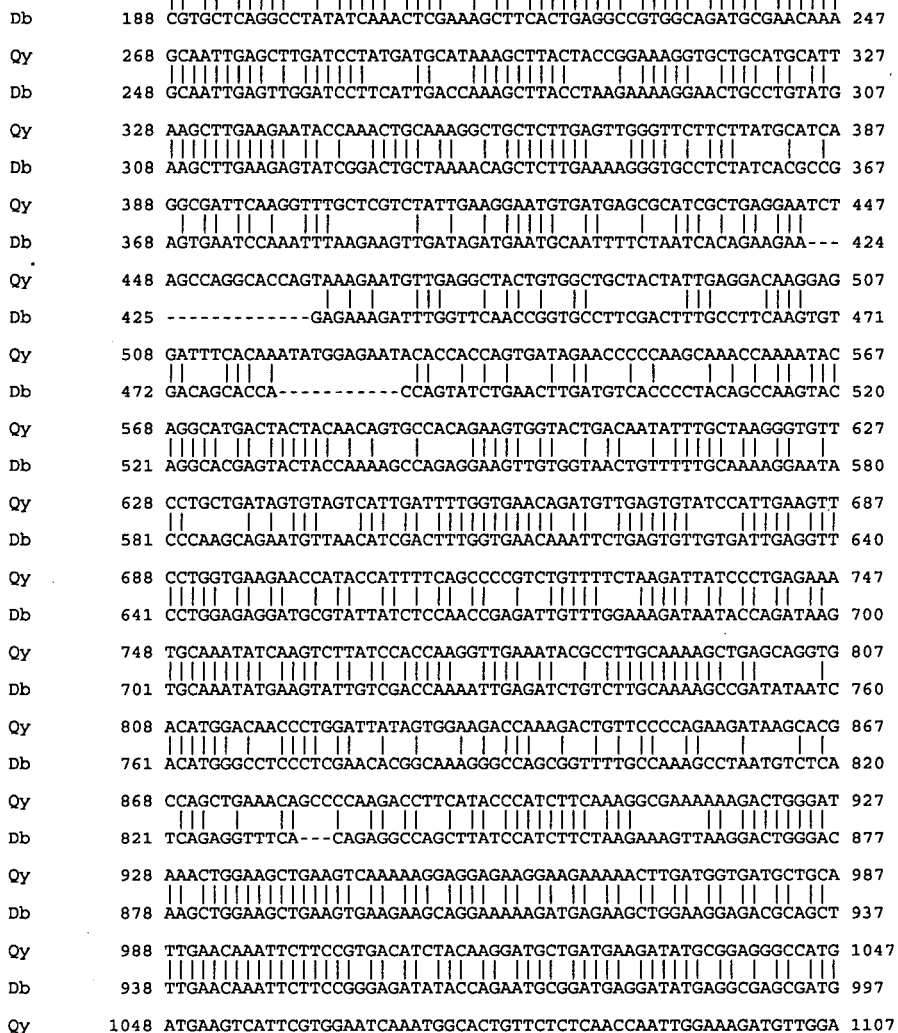
```
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             28.8%;  Score 418.4;  DB 3;  Length 1361;
  Best Local Similarity   62.9%;  Pred. No. 8.2e-82;
  Matches  682;  Conservative    0;  Mismatches  396;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db         54 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 113

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db        114 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 173

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        174 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 233

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        234 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 293

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        294 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 353

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        354 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 413

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        414 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 470

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        471 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 530

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        531 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 590

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        591 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 650

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        651 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 710

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        711 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 770

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        771 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 830

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        831 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 887

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        888 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 947

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db        948 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1007

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db       1008 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1067

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
Db       1068 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1127

Qy       1168 AGTT 1171
```

```
Db       1128 TCTT 1131

RESULT 7
ADG87792
ID   ADG87792 standard; cDNA; 1077 BP.
XX
AC   ADG87792;
XX
DT   22-APR-2004  (first entry)
XX
DE   A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #234.
XX
KW   Pathogen infection-related gene; plant; Peronospora parasitica;
KW   defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW   oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200222675-A2.
XX
PD   21-MAR-2002.
XX
PF   14-SEP-2001; 2001WO-US028506.
XX
PR   15-SEP-2000; 2000US-0232778P.
PR   22-JUN-2001; 2001US-0300183P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (UYNC-) UNIV NORTH CAROLINA.
PA   (GLAZ/) GLAZEBROOK J.
PA   (WANG/) WANG X.
PA   (DANG/) DANGL J L.
PA   (EULG/) EULGEM T.
PA   (ZHUT/) ZHU T.
XX
PI   Glazebrook J,  Wang X,  Dangl JL,  Eulgem T,  Zhu T;
XX
DR   WPI; 2002-292409/33.
XX
PT   Novel isolated polynucleotide, useful for conveying pathogen resistance
PT   to plants, and for identifying plants infected with a pathogen.
XX
PS   Claim 3; SEQ ID NO 234; 605pp; English.
XX
CC   The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC   ADG87557}) whose expression is altered in response to pathogen infection,
CC   and to homologues of these genes from other plants or fungi, especially
CC   from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC   cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC   expression of genes of the invention was upregulated or downregulated in
CC   Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC   indicating that they play a role in defence mechanisms. The genes of the
CC   invention are regulated by RPP7 or RRP8 which act via unconventional
CC   signalling cascades, or by the RPP4-dependent pathway. The invention also
CC   relates to polypeptides encoded by the pathogen infection-related genes;
CC   promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC   ; expression cassettes, host cells and pathogen-resistant transgenic
CC   plants and their progeny comprising a polynucleotide of the invention;
CC   and a method of identifying a plant cell infected with a pathogen. The
CC   polynucleotide sequences and methods of the invention are useful for
CC   identifying plants infected with a pathogen, and for conferring
CC   resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC   neamotodes and insects (e.g., aphids). The present sequence represents an
CC   Arabidopsis thaliana gene whose expression is altered in response to
CC   Peronospora parasitica infection. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1077 BP; 337 A; 219 C; 277 G; 244 T; 0 U; 0 Other;

  Query Match             28.7%;  Score 417.6;  DB 6;  Length 1077;
  Best Local Similarity   63.0%;  Pred. No. 1.1e-81;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
```

```
                  | ||||| || ||  | |  || || ||||| |||||||||||||||||||| || | |
Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077

RESULT 8
AAC42413
ID   AAC42413 standard; DNA; 1117 BP.
XX
AC   AAC42413;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 35460.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
```

PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.0%; Score 393.2; DB 3; Length 1117;
Best Local Similarity 62.7%; Pred. No. 2.6e-76;
Matches 677; Conservative 0; Mismatches 373; Indels 30; Gaps 3;



RESULT 9
ADF51380
ID ADF51380 standard; cDNA; 388 BP.
XX
AC ADF51380;
XX
DT 12-FEB-2004 (first entry)
XX

```
DE   Maize Rar interactior 2 (RarInt2) DNA seq id 3.
XX
KW   plant protectant; gene therapy; Rar1; disease resistance; plant;
KW   immunogen; maize; Rar interactor 2; RarInt2; gene; ss.
XX
OS   Zea mays.
XX
PN   US2003167504-A1.
XX
PD   04-SEP-2003.
XX
PF   12-FEB-2002; 2002US-00074473.
XX
PR   12-FEB-2001; 2001US-0268157P.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Crane EH;
XX
DR   WPI; 2003-898130/82.
XX
PT   New maize Rar1-interactor nucleic acid useful for enhancing disease
PT   resistance in plants, as a probe or amplification primer, in recombinant
PT   expression of polypeptides, or as an immunogen in the preparation or
PT   screening of antibodies.
XX
PS   Claim 1; SEQ ID NO 3; 30pp; English.
XX
CC   The invention describes an isolated nucleic acid capable of interacting
CC   with Rar1. The nucleic acid comprises a member selected from: (a) a
CC   polynucleotide comprising a sequence of 388 (S1), 571 (S2) or 691 (S3) bp
CC   fully defined in the specification; (b) a polynucleotide amplified from a
CC   Zea mays nucleic acid library using primers that selectively hybridise,
CC   under high stringency conditions, to loci within the polynucleotide in
CC   (a); (c) a polynucleotide encoding a maize Rar1-interactor protein; (d) a
CC   polynucleotide having at least 80 or 90% sequence identity to (a); (e) a
CC   polynucleotide which hybridises under high stringency conditions to (a);
CC   and (f) a polynucleotide complementary to any of the above
CC   polynucleotides. The nucleic acid molecule and methods are useful in
CC   enhancing disease resistance in plants. The nucleic acid molecule may
CC   also be used as a probe or an amplification primer, in recombinant
CC   expression of polypeptides, or as an immunogen in the preparation and/or
CC   screening of antibodies. This sequence represents a maize Rar interactor
CC   2 (RarInt2) polynucleotide.
XX
SQ   Sequence 388 BP; 102 A; 76 C; 93 G; 117 T; 0 U; 0 Other;

  Query Match            26.7%;  Score 388;  DB 10;  Length 388;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-75;
  Matches  388;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1043 CCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATG 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATG 60

Qy       1103 TTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAAT 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAAT 120

Qy       1163 ACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTC 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTTGGAAACTATGACCTAATTC 180

Qy       1223 TGCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGT 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCCCACCCATAGTGCCATGAGCTTGCTTGGTTAAGTCTCTGCTTTTGTAAGCTTTCTGT 240

Qy       1283 ATGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATG 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGACAGTCTAGCGTGTGTCTGAGTGCTACTTGCCGTAGCTTTTGATGGACATAGGTATG 300

Qy       1343 CTATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTATTATGTGGTATCGACCCCGAACCATCTGTGGTATCAAAGACAGCAAACTATAATGTA 360

Qy       1403 TGGAATGCCTATTCTTTTTGGTCTAAAA 1430
             ||||||||||||||||||||||||||||
Db        361 TGGAATGCCTATTCTTTTTGGTCTAAAA 388


RESULT 10
ABL71197
ID   ABL71197 standard; cDNA; 287 BP.
XX
AC   ABL71197;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:571.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 571; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC   hybridisation or amplification technologies, as query sequences to
CC   determine homology of known sequences, as probe for use in Southern or
CC   Northern hybridisation, and to identify the presence of and/or to
CC   determine the degree of similarity between two (or more) nucleic acid
CC   sequences
XX
SQ   Sequence 287 BP; 89 A; 60 C; 70 G; 68 T; 0 U; 0 Other;

  Query Match            19.7%;  Score 286;  DB 6;  Length 287;
  Best Local Similarity  100.0%;  Pred. No. 6e-53;
  Matches  286;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        297 AGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGC 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGC 60

Qy        357 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAA 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAA 120

Qy        417 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGC 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGC 180

Qy        477 TACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACC 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACC 240

Qy        537 AGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 286
```

RESULT 11
ACL25190/c
ID ACL25190 standard; DNA; 581 BP.
XX
AC ACL25190;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #15181.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 581 BP; 157 A; 151 C; 101 G; 172 T; 0 U; 0 Other;

```
  Query Match            16.5%;  Score 240.2;  DB 9;  Length 581;
  Best Local Similarity  79.6%;  Pred. No. 8.4e-43;
  Matches  284;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy        847 GTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCA 906
              | ||| ||||||||||      |||||||| ||||||  || || || || || || |||
Db        569 GCTCCTCAGAAGATAAATGTACCAGCTGAATCAGCCCAGAGGCCATCTTATCCTTCATCA 510

Qy        907 AAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAA 966
              ||  | ||||| |||||||||| || || |||||||| |||||  |||||||||| ||
Db        509 AAATCCAAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAACAGGAGAAGGATGAG 450

Qy        967 AAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGAT 1026
              |||||||| |||||||||||||||||| || |||||||| || ||||  |||||||||
Db        449 AAACTTGACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATTTACAGTGATGCTGAT 390

Qy       1027 GAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCA 1086
              |||||||||| || || |||||||||| || ||||| || ||||| || |||||||||
Db        389 GAAGATATGCGTAGAGCAATGATGAAGTCCTTTGTGGAGTCTAATGGAACCGTTCTCTCA 330

Qy       1087 ACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAG 1146
              ||||| |||||||||| ||   ||||| ||| ||||| ||||| |||||||| ||||||
Db        329 ACCAACTGGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCTCCTGATGGAATGGAG 270

Qy       1147 CTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTAAATCCAGGTCTT 1203
              |||||||||||||| || ||| |    |   |    | |||||||||| |||| |
Db        269 CTCAAGAAGTGGGAGTATTAATGAAACCAGATCATCCATGTTGTAAATCCTGGTCGT 213
```

RESULT 12
ACN47711
ID ACN47711 standard; cDNA; 585 BP.
XX
AC ACN47711;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton primed seed EST Clone ID: LIB3825-014-Q1-K6-D9, SEQ:2492.
XX
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
KW variety DP50B; library LIB3825; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS Gossypium hirsutum.
XX
PN US2004123340-A1.
XX
PD 24-JUN-2004.
XX
PF 12-DEC-2001; 2001US-00021323.
XX
PR 14-DEC-2000; 2000US-0255619P.
XX
PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX
PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
XX
DR WPI; 2004-479808/45.
XX
PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX

```
PS   Claim 1; SEQ ID NO 2492; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 585 BP; 186 A; 117 C; 134 G; 148 T; 0 U; 0 Other;

  Query Match             15.2%;  Score 221.8;  DB 13;  Length 585;
  Best Local Similarity   62.7%;  Pred. No. 9.2e-39;
  Matches  362;  Conservative    0;  Mismatches  212;  Indels    3;  Gaps    1;

Qy        229 CTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATG 288
              ||| |||   | |||| |||||||| || |||||||||||||||||| | |||||
Db          1 CTCAACAATCTCCCTGACGCTGTGGCAGAAGCTAACAAAGCAATTGAGTTGGATCCGTCC 60

Qy        289 ATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACT 348
              ||| |||| | ||||  || |||| | |||| || || |||||||| || || ||||||
Db         61 ATGTCTAAATCCTACTTGCGTAAAGCTACTGCGTGTATGAAGCTTGAGGAGTATCAAACT 120

Qy        349 GCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGT 408
              || |||||||  | |||   ||  || |||  ||| |||  || || || ||| |
Db        121 GCTAAGGCTGCGTTGGAGACTGGGGCTGCTTTGGCACCAGAAGACTCGAGATTTTCCAAG 180

Qy        409 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAT 468
              |  | || |||||||| |||||||| || ||||||  | |  ||  ||||    |||
Db        181 TTGATTAAAGAATGTGAAGAGCGCATTGCAGAGGAAAATGGTGAGTTACCAAAGGAGACG 240

Qy        469 GTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAAT 528
              | || |         |  |  | |    |  || ||    |   |      ||| |
Db        241 TTGGAAGAGGTGCCAACAAATGTTGTACCTGCGAAAGACATCTCTTCTGTCCAGGATATA 300

Qy        529 ACACCACCAGTGA---TAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAAC 585
              |    ||  |||   | | | | |||  |||| |||  ||||||||||| |  ||  |
Db        301 CCTGATCCCATGACTGTTGCAGCACCACCCAAATCAACTTACAGGCATGAATTTTATCAG 360

Qy        586 AGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTC 645
              |   |    ||||||||  | ||||||||||| |||||  | ||    ||  ||||
Db        361 AAACCAGAGGAAGTGGTTGTCACAATATTTGCCAAGGGAATACCACGCGAGTGTGTTAAA 420

Qy        646 ATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATAC 705
              || ||| |||||||||| ||  | |||||  |  |  |  ||  ||  ||||  | ||
Db        421 GTTAATTATGGTGAACAAATACTAAGTGTTGCTGTCAATGCACCCGGCAAAGATGCTTAT 480

Qy        706 CATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTA 765
              ||| | || || ||  | |||   ||||| || ||||| || |||| |||| | || ||
Db        481 CATCTCCAACCTCGCTTATTTGGAAAGATAATACCTGACAAGTGCAGATATGATGTTTTG 540

Qy        766 TCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGC 802
              || |||||  |||||||  |  | ||||||||||| |
Db        541 TCAACCAAAATTGAAATTAGGTTAGCAAAAGCTGAAC 577


RESULT 13
ACN59636
ID   ACN59636 standard; cDNA; 534 BP.
XX
AC   ACN59636;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Cotton gynoecium tissue EST Clone ID: LIB3829-033-Q1-K6-G10, SEQ:14417.
XX
KW   Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW   variety Nucotton33B; library LIB3829; molecular tag; molecular marker;
KW   genetic mapping; molecular mapping; seed germination; plant growth;
KW   plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS   Gossypium hirsutum.
XX
PN   US2004123340-A1.
XX
PD   24-JUN-2004.
XX
PF   12-DEC-2001; 2001US-00021323.
XX
PR   14-DEC-2000; 2000US-0255619P.
XX
PA   (DEIK/) DEIKMAN J.
PA   (FENG/) FENG P C C.
PA   (FINC/) FINCHER K L.
PA   (ZIEG/) ZIEGLER T E.
XX
PI   Deikman J,  Feng PCC,  Fincher KL,  Ziegler TE;
XX
DR   WPI; 2004-479808/45.
XX
PT   New isolated nucleic acid molecule that encodes a plant protein or its
PT   fragment, useful for isolating a variety of agronomically significant
PT   genes associated with plant growth, quality or yield, and as molecular
PT   tags to map genes.
XX
PS   Claim 1; SEQ ID NO 14417; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
```

```
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the US
CC   patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 534 BP; 181 A; 105 C; 115 G; 133 T; 0 U; 0 Other;

  Query Match             15.0%;  Score 217.8;  DB 13;  Length 534;
  Best Local Similarity   65.2%;  Pred. No. 6.8e-38;
  Matches  337;  Conservative    0;  Mismatches  177;  Indels    3;  Gaps    1;

Qy        542 TAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGG 601
              | | | | |||  |||| |||  ||||||||||| |  ||  | |    |     |||||||
Db         17 TTGCAGCACCACCCAAATCAACTTACAGGCATGAATTTTATCAGAAACCAGAGGAAGTGG 76

Qy        602 TACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAAC 661
              |  | ||||||||||| |||||  | ||    ||  ||||      || ||| ||||||||
Db         77 TTGTCACAATATTTGCCAAGGGAATACCACGCGAGTGTGTTAAAGTTAATTATGGTGAAC 136

Qy        662 AGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTC 721
              | ||  | |||||  |  |  | |  || ||  ||||  | || ||| | || || ||
Db        137 AAATACTAAGTGTTGCTGTCAATGCACCCGGCAAAGATGCTTATCATCTCCAACCTCGCT 196

Qy        722 TGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAA 781
              | |||   ||||| || ||||| || |||| |||| | || || || |||||  ||||||
Db        197 TATTTGGAAAGATAATACCTGACAAGTGCAGATATGATGTTTTGTCAACCAAAATTGAAA 256

Qy        782 TACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAA 841
              |  |  | ||||||||||| |    |   ||||||| |||| || | |||       |
Db        257 TTAGGTTAGCAAAAGCTGAACCAATTCAATGGACATCCCTCGAATTTAGCATGGAAGTTG 316

Qy        842 AGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCAT 901
                    |   ||  |    |  | | ||   ||   | |||||| || |||||||
Db        317 CTGTGCCCCAAAGGATTAATGTATCTTCTGTTGCTGCAAATCAAAGACCATCGTACCCAT 376

Qy        902 CTTCAAAGGCGAAAA---AAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGA 958
              | || ||  | ||||     || |||||||||| | |||||| |||| ||||||||||||||
Db        377 CCTCCAAACCAAAAAGGGTTGATTGGGATAAAATCGAAGCTCAAGTGAAAAAGGAGGAGA 436

Qy        959 AGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGG 1018
              | ||  |||||  || |||||||||||| || |||||  |||||  |||||  |||||  ||   |||
Db        437 AAGATGAAAAGCTAGATGGTGATGCAGCTTTGAATAAATTTTTCCGCGACATTTATCAGG 496

Qy       1019 ATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTC 1055
              |||||||  ||  |||     |  |||||||||   |||  ||
Db        497 ATGCTGACGAGGATGCAAGAAGGGCCATGCAGAAATC 533


RESULT 14
ABL75330
ID   ABL75330 standard; cDNA; 272 BP.
XX
AC   ABL75330;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:4704.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 4704; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC   hybridisation or amplification technologies, as query sequences to
CC   determine homology of known sequences, as probe for use in Southern or
CC   Northern hybridisation, and to identify the presence of and/or to
CC   determine the degree of similarity between two (or more) nucleic acid
CC   sequences
XX
SQ   Sequence 272 BP; 95 A; 54 C; 70 G; 49 T; 0 U; 4 Other;

  Query Match             14.9%;  Score 216.4;  DB 6;  Length 272;
  Best Local Similarity   90.9%;  Pred. No. 1.1e-37;
  Matches  241;  Conservative    0;  Mismatches   19;  Indels    5;  Gaps    1;

Qy        784 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 60

Qy        844 ACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCT 903
                || |||||||||||||||||||||||| |||     | || ||||||||||||||||
Db         61 GCTATTCCCCAGAAGATAAGCACGCCAGCNGAA-----CACANGACCTTCATACCCATCT 115

Qy        904 TCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAA 963
              |||||| | ||||| |||||||||||||||||||||||||||| ||||||||||||||||
Db        116 TCAAAGTCAAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAA 175

Qy        964 GAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCT 1023
              |||||||||||| |||||||||||||||||||| |||||||||||  ||||||||||||||
Db        176 GAAAAACTTGAAGGTGATGCTGCATTGAACAAGTTCTTCCGTGATCTCTACAAGGATGCT 235

Qy       1024 GATGAAGATATGCGGAGGGCCATGA 1048
              |||||||||||  |  ||||  ||
Db        236 GATGAAGATATGGCGGNGGCCTGGA 260


RESULT 15
ACL25189/c
ID   ACL25189 standard; DNA; 541 BP.
XX
AC   ACL25189;
XX
DT   27-OCT-2003  (revised)
DT   17-OCT-2003  (first entry)
XX
```

DE DNA clone originating in barley containing SNP encoding sequence #15180.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 541 BP; 151 A; 144 C; 89 G; 157 T; 0 U; 0 Other;

Query Match 14.7%; Score 214.4; DB 9; Length 541;
Best Local Similarity 80.4%; Pred. No. 3.8e-37;
Matches 251; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```
Qy        892 TCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAG 951
              || || || || |||||  | ||||| |||||||||||| || || |||||||| |||||
Db        539 TCTTATCCTTCATCAAAATCCAAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAA 480

Qy        952 GAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATC 1011
               |||||||||| || |||||||| |||||||||||||||||||| || |||||||| ||
Db        479 CAGGAGAAGGATGAGAAACTTGACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATT 420

Qy       1012 TACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAAT 1071
              ||||  |||||||||||||||||||| || || |||||||||| || ||||| || |||
Db        419 TACAGTGATGCTGATGAAGATATGCGTAGAGCAATGATGAAGTCCTTTGTGGAGTCTAAT 360

Qy       1072 GGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 1131
              || || |||||||||||||| |||||||||| ||   ||||| ||| ||||| |||||
Db        359 GGAACCGTTCTCTCAACCAACTGGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCT 300

Qy       1132 CCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTTTGTA 1191
              |||||||| |||||||||||||||||||| || ||| |     |   |     | |||||
Db        299 CCTGATGGAATGGAGCTCAAGAAGTGGGAGTATTAATGAAACCAGATCATCCATGTTGTA 240

Qy       1192 AATCCAGGTCTT 1203
              ||||| |||| |
```

Db 239 AATCCTGGTCGT 228

Search completed: April 23, 2005, 03:10:58
Job time : 847.735 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:46:01 ; Search time 260.537 Seconds
(without alignments)
9137.976 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcggcacgagcggcac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	164.4	11.3	272	4	US-09-313-294A-4573	Sequence 4573, Ap
2	156.4	10.7	2053	3	US-09-227-357-45	Sequence 45, Appl
3	76.2	5.2	198	4	US-09-248-796A-2330	Sequence 2330, Ap
4	74.2	5.1	1735	4	US-09-949-016-1499	Sequence 1499, Ap
5	74.2	5.1	1982	4	US-09-016-434-1067	Sequence 1067, Ap
6	74.2	5.1	1982	4	US-09-825-497A-40	Sequence 40, Appl
7	74.2	5.1	1987	4	US-09-825-497A-39	Sequence 39, Appl
8	74.2	5.1	1987	4	US-09-517-779-1	Sequence 1, Appli
9	73.2	5.0	425	4	US-09-513-999C-4027	Sequence 4027, Ap
10	63.4	4.4	47493	4	US-09-949-016-13241	Sequence 13241, A
11	62.4	4.3	2735	4	US-09-551-974A-101	Sequence 101, App
12	62.4	4.3	2735	4	US-09-565-501A-101	Sequence 101, App
13	62.4	4.3	2735	4	US-09-639-206A-101	Sequence 101, App
14	62.4	4.3	2735	4	US-09-874-923-101	Sequence 101, App
15	62.4	4.3	3012	4	US-09-551-974A-94	Sequence 94, Appl
16	62.4	4.3	3012	4	US-09-565-501A-94	Sequence 94, Appl
17	62.4	4.3	3012	4	US-09-639-206A-94	Sequence 94, Appl
18	62.4	4.3	3012	4	US-09-874-923-94	Sequence 94, Appl
19	62.4	4.3	3134	2	US-08-533-669A-1	Sequence 1, Appli
20	62.4	4.3	3134	3	US-09-183-861-1	Sequence 1, Appli
21	62.4	4.3	3134	3	US-09-022-765-1	Sequence 1, Appli
22	62.4	4.3	3134	4	US-09-551-974A-1	Sequence 1, Appli
23	62.4	4.3	3134	4	US-09-565-501A-1	Sequence 1, Appli
24	62.4	4.3	3134	4	US-09-639-206A-1	Sequence 1, Appli
25	62.4	4.3	3134	4	US-09-874-923-1	Sequence 1, Appli
26	62.4	4.3	3134	4	US-08-798-841-1	Sequence 1, Appli
27	62.4	4.3	4233	4	US-09-551-974A-99	Sequence 99, Appl
28	62.4	4.3	4233	4	US-09-565-501A-99	Sequence 99, Appl
29	62.4	4.3	4233	4	US-09-639-206A-99	Sequence 99, Appl
30	62.4	4.3	4233	4	US-09-874-923-99	Sequence 99, Appl
31	62.4	4.3	4917	4	US-09-551-974A-100	Sequence 100, App
32	62.4	4.3	4917	4	US-09-565-501A-100	Sequence 100, App
33	62.4	4.3	4917	4	US-09-639-206A-100	Sequence 100, App
34	62.4	4.3	4917	4	US-09-874-923-100	Sequence 100, App
35	62.4	4.3	4929	4	US-09-551-974A-98	Sequence 98, Appl
36	62.4	4.3	4929	4	US-09-565-501A-98	Sequence 98, Appl
37	62.4	4.3	4929	4	US-09-639-206A-98	Sequence 98, Appl
38	62.4	4.3	4929	4	US-09-874-923-98	Sequence 98, Appl
39	62.2	4.3	989	4	US-09-270-767-12741	Sequence 12741, A
40	59.8	4.1	2221	3	US-09-301-978C-1	Sequence 1, Appli
41	59.8	4.1	2229	4	US-09-949-016-1647	Sequence 1647, Ap
c 42	59.8	4.1	2274	4	US-09-220-132-188	Sequence 188, App
c 43	58.4	4.0	7218	1	US-08-232-463-14	Sequence 14, Appl
44	54.4	3.7	1304	4	US-09-902-540-2596	Sequence 2596, Ap
45	54.4	3.7	16584	4	US-09-902-540-1119	Sequence 1119, Ap

ALIGNMENTS

RESULT 1
US-09-313-294A-4573
; Sequence 4573, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4573
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348654H1
; NAME/KEY: unsure
; LOCATION: 92, 100, 103, 105, 112, 114, 140, 204, 210, 244-245, 250, 257, 261-262,
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4573

Query Match 11.3%; Score 164.4; DB 4; Length 272;
Best Local Similarity 77.3%; Pred. No. 2.2e-31;
Matches 214; Conservative 0; Mismatches 58; Indels 5; Gaps 2;

```
Qy      407 GTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGA 466
            ||||||||||||||||||||||||| |||| ||||||||||||||||||||||||| |||||
Db        1 GTCTATTGAAGGAATGTGATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGA 60

Qy      467 ATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGA 526
            |||| |||||| |||| ||||||||| |||| ||||||| || | || ||| | |||| |
Db       61 ATGTCGAGGCTCCTGTAGCTGCTACTGTTGANGACAAGGNGGNTNTCGCAANTNTGGATA 120

Qy      527 ATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACA 586
            ||||||| |||||| |||| |||||||||||||| ||||| ||||||||||||||||||
Db      121 ATACACCGCCAGTGGTAGANCCCCCAAGCAAACCTAAATATAGGCATGACTACTACAACA 180

Qy      587 GTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCA 646
            |||||||||||  |       ||     ||||  | ||   | | ||||||| ||||
Db      181 GTGCCACAGAATGG---TCCTCACATNTTGCCNAAGGGTGTTCCGCTGATAATGTA--TC 235

Qy      647 TTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 683
              |   ||  |||  |||||| |||     || || |
Db      236 ATTGATTCNNTGATNAGATGTNGAGNNCNNCCCTTCA 272
```

```
RESULT 2
US-09-227-357-45
; Sequence 45, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P1
;  CURRENT APPLICATION NUMBER: US/09/227,357
;  CURRENT FILING DATE: 1999-01-08
;  EARLIER APPLICATION NUMBER: PCT/US98/13684
;  EARLIER FILING DATE: 1998-07-07
;  EARLIER APPLICATION NUMBER: 60/051,926
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,793
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,925
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,929
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,803
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,732
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,931
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,932
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,916
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,930
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,918
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,920
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,733
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,795
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,919
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,928
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/055,722
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,723
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,948
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,949
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,953
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,950
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,947
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,964
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/056,360
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,684
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,984
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,954
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/058,785
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,664
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,660
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,661
;  EARLIER FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 2053
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-227-357-45

  Query Match           10.7%;  Score 156.4;  DB 3;  Length 2053;
  Best Local Similarity  56.3%;  Pred. No. 5.7e-29;
  Matches  339;  Conservative    0;  Mismatches  251;  Indels   12;  Gaps    2;

Qy       560 CAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTA 619
            |||||  || | ||||||  ||  | |   |  |||| ||  | ||| | |   |
Db      1416 CAAAAATCAAGTATGACTGGTATCAAACAGAATCTCAAGTAGTCATTACACTTATGATCA 1475

Qy       620 AGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCA 679
            ||  |||||        ||  ||||    | || |||   ||| |   |||| ||  |
Db      1476 AGAATGTTCAGAAGAATGATGTAAATGTGGAATTTTCAGAAAAAGAGTTGTCTGCTTTGG 1535

Qy       680 TTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCC 739
            || || |||||  || |||     ||| ||||  | |       ||  ||  |   || || |
Db      1536 TTAAACTTCCTTCTGGAGAGGATTACAATTTGAAACTGGAACTTCTTCATCCTATAATAC 1595

Qy       740 CTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTG 799
            | ||  |  |||  | | ||||  | || || ||| |||||||     ||   |||  | |
Db      1596 CAGAACAGAGCACGTTTAAAGTACTTTCAACAAAGATTGAAATTAAACTGAAAAAGCCAG 1655

Qy       800 AGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGA 859
            ||   |||| ||||   ||  || ||         ||   |  | |       | | |  ||
Db      1656 AGGCTGTGAGATGGGAAAAGCTAGA------GGGGCAAGGAGATGTGCCTACGCCAAAAC 1709

Qy       860 TAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAG 919
            |  |    ||| || ||    || |   || ||||   |  |||         | || |
Db      1710 AATTCGTAGCAGATGTAAAGAACCTATATCCATCATCATCTCCTT------ATACAAGAA 1763

Qy       920 ACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTG 979
            | |||||||||  |||  | |||  ||||| | || || ||| | |||||  | || || |
Db      1764 ATTGGGATAAATTGGTTGGTGAGATCAAAGAAGAAGAAAAGAATGAAAAGTTGGAGGGAG 1823

Qy       980 ATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGA 1039
            ||||  || || |||| ||| || |   | |||||    |||| |  ||| ||  ||
Db      1824 ATGCAGCTTTAAACAGATTATTTCAGCAGATCTATTCAGATGGTTCTGATGAAGTGAAAC 1883

Qy      1040 GGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAG 1099
            | |||||||  || || ||  |||| ||   ||| || ||| |    |||| ||| |
Db      1884 GTGCCATGAACAAATCCTTTATGGAGTCGGGTGGTACAGTTTTGAGTACCAACTGGTCTG 1943

Qy      1100 ATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGG 1159
            ||||  ||   || ||| || |||   |  || ||||||| ||||||     || ||||
Db      1944 ATGTAGGTAAAAGGAAAGTTGAAATCAATCCTCCTGATGATATGGAATGGAAAAAGTACT 2003

Qy      1160 AA 1161
            ||
Db      2004 AA 2005


RESULT 3
US-09-248-796A-2330
; Sequence 2330, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
```

```
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2330
;    LENGTH: 198
;    TYPE: DNA
;    ORGANISM: Candida albicans
US-09-248-796A-2330

  Query Match             5.2%;  Score 76.2;  DB 4;  Length 198;
  Best Local Similarity   64.4%;  Pred. No. 2.2e-09;
  Matches  114;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy        991 AACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATG 1050
              ||  | || |||||  | || | ||| |||| ||| ||||||    | || || ||||||
Db         22 AATGACTTTTTCCGGAAAATATTCAAAGATGTTGACGAAGATTCTAGAAGAGCGATGATG 81

Qy       1051 AAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCA 1110
              |||   |  ||  |||| ||||| |||||| |  | || | |||| | || | |  ||
Db         82 AAGAGTTATGTTCAATCGAATGGAACTGTTTTGACTACAAGTTGGGATGAAGCTAAAGAT 141

Qy       1111 AAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              ||  |  | || |    || || ||||| |||||  ||||||| |||||    |||
Db        142 AAAGAATTTGAGGTTTTACCACCAGATGGGATGGAAGTCAAGAAATGGGATACGTAA 198


RESULT 4
US-09-949-016-1499
; Sequence 1499, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1499
;    LENGTH: 1735
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-1499

  Query Match             5.1%;  Score 74.2;  DB 4;  Length 1735;
  Best Local Similarity   55.1%;  Pred. No. 1.9e-08;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         82 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 141

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        142 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 201

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |       |   |||     | ||  |||  || || |
Db        202 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 261

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||||| ||    | |       || | ||| |||||    | |||      |||
Db        262 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 321

Qy        328 AAGCTTGAAGAATACCAAACTGC 350
              |||| |   | | ||   | ||
Db        322 GCACTGGGCAAGTTCCGGGCCGC 344


RESULT 5
US-09-016-434-1067
; Sequence 1067, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1067:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1982 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1122930
US-09-016-434-1067

  Query Match             5.1%;  Score 74.2;  DB 4;  Length 1982;
  Best Local Similarity   55.1%;  Pred. No. 2e-08;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         60 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 119

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        120 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 179

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |       |   |||     | ||  |||  || || |
Db        180 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 239

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||||| ||    | |       || | ||| |||||    | |||      |||
Db        240 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 299

Qy        328 AAGCTTGAAGAATACCAAACTGC 350
```

```
                  || |   | | ||   | ||
Db        300 GCACTGGGCAAGTTCCGGGCCGC 322


RESULT 6
US-09-825-497A-40
; Sequence 40, Application US/09825497A
; Patent No. 6599742
; GENERAL INFORMATION:
;  APPLICANT: Honkanen, Richard  E.
;  APPLICANT:  Dean, Nicholas  M.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODUATION OF HUMAN SERINE/THREONINE PRO
;  TITLE OF INVENTION:  PHOSPHATASE GENE EXPRESSION
;  FILE REFERENCE: ISPH-0572
;  CURRENT APPLICATION NUMBER: US/09/825,497A
;  CURRENT FILING DATE:  2001-04-06
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
;   LENGTH: 1982
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-497A-40

  Query Match             5.1%;  Score 74.2;  DB 4;  Length 1982;
  Best Local Similarity   55.1%;  Pred. No. 2e-08;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         60 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 119

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        120 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 179

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |     |  |||     | ||  |||  || || |
Db        180 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 239

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||||| ||    | |      || | ||| |||||    | |||      |||
Db        240 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 299

Qy        328 AAGCTTGAAGAATACCAAACTGC 350
              || |   | | ||   | ||
Db        300 GCACTGGGCAAGTTCCGGGCCGC 322


RESULT 7
US-09-825-497A-39
; Sequence 39, Application US/09825497A
; Patent No. 6599742
; GENERAL INFORMATION:
;  APPLICANT: Honkanen, Richard  E.
;  APPLICANT:  Dean, Nicholas  M.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODUATION OF HUMAN SERINE/THREONINE PRO
;  TITLE OF INVENTION:  PHOSPHATASE GENE EXPRESSION
;  FILE REFERENCE: ISPH-0572
;  CURRENT APPLICATION NUMBER: US/09/825,497A
;  CURRENT FILING DATE:  2001-04-06
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-497A-39

  Query Match             5.1%;  Score 74.2;  DB 4;  Length 1987;
  Best Local Similarity   55.1%;  Pred. No. 2e-08;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         65 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 124

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        125 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 184

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |     |  |||     | ||  |||  || || |
Db        185 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 244

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||||| ||    | |      || | ||| |||||    | |||      |||
Db        245 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 304

Qy        328 AAGCTTGAAGAATACCAAACTGC 350
              || |   | | ||   | ||
Db        305 GCACTGGGCAAGTTCCGGGCCGC 327


RESULT 8
US-09-517-779-1
; Sequence 1, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
;  APPLICANT:  Luo, Ying
;  APPLICANT:  Xu, Xiang
;  TITLE OF INVENTION:  Cell Cycle Proteins Associated with Rad9, Compositions
;  TITLE OF INVENTION:  and Methods Of Use
;  FILE REFERENCE:  A68293/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/517,779
;  CURRENT FILING DATE:  2000-03-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-517-779-1

  Query Match             5.1%;  Score 74.2;  DB 4;  Length 1987;
  Best Local Similarity   55.1%;  Pred. No. 2e-08;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         65 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 124

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        125 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 184

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |     |  |||     | ||  |||  || || |
Db        185 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 244

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||||| ||    | |      || | ||| |||||    | |||      |||
Db        245 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 304

Qy        328 AAGCTTGAAGAATACCAAACTGC 350
              || |   | | ||   | ||
Db        305 GCACTGGGCAAGTTCCGGGCCGC 327


RESULT 9
US-09-513-999C-4027
; Sequence 4027, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
```

```
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 4027
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 46..417
US-09-513-999C-4027

  Query Match             5.0%;  Score 73.2;  DB 4;  Length 425;
  Best Local Similarity   57.4%;  Pred. No. 1.8e-08;
  Matches  132;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db        127 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 186

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
              | | || ||||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        187 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 246

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              |||  || |||  || |       |    |||     | ||  |||  || || |
Db        247 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 306

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGC 317
              || |||||||| ||    | |       || | ||| |||||    | |||
Db        307 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGC 356


RESULT 10
US-09-949-016-13241
; Sequence 13241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13241
;   LENGTH: 47493
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13241

  Query Match             4.4%;  Score 63.4;  DB 4;  Length 47493;
  Best Local Similarity   54.5%;  Pred. No. 4.6e-05;
  Matches  127;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy        118 GCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCC 177
              | || ||  | | | |||| ||||   |||| | || ||||||||||||||||||||
Db       8641 GTCTCCGCAGCCAAGGACTACGAGAACGCCATCAAGTTCTACAGCCAGGCCATCGAGCTG 8700

Qy        178 GGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAAC 237
              |||  ||  |||  |  ||||| | |||||  || |||  || |        |    |
Db       8701 AACCCCAGCAATGCCATCTACTATGGCAACCGCAGCCTGGCCTACCTGCGCACTGAGTGC 8760

Qy        238 TACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAA 297
              ||     | ||  |||  || || |     || ||||||| ||    | |       ||
Db       8761 TATGGCTACGCGCTGGGAGACGCCACGCGGGCCATTGAGCTGGACAAGAAGTACATCAAG 8820

Qy        298 GCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGC 350
              | ||| |||||    | |||       |||    || |   | | ||   | ||
Db       8821 GGTTATTACCGCCGGGCTGCCAGCAACATGGCACTGGGCAAGTTCCGGGCCGC 8873


RESULT 11
US-09-551-974A-101
; Sequence 101, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C5
;  CURRENT APPLICATION NUMBER: US/09/551,974A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-551-974A-101

  Query Match             4.3%;  Score 62.4;  DB 4;  Length 2735;
  Best Local Similarity   51.4%;  Pred. No. 2.1e-05;
  Matches  144;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         65 AAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCG 124
              |||||  |  |   |  | |||| | |   | |   | | |  ||  ||  ||||
Db       1676 AAGCCTACATCGATCCTGAGATCGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCA 1735

Qy        125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCG 184
              || ||  | |||     ||||   |     ||||   ||||||||  |   |   || |
Db       1736 AGGAGGATAAGTTCCCCGAGGCCGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTG 1795

Qy        185 CCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTG 244
              ||     |  | |||   | | |||||     ||| ||||||||| ||   || ||  |
Db       1796 CCGAGCACACCTCCTACAGCAATCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACG 1855

Qy        245 AGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACT 304
              | ||  |  ||| || | ||    |||||||  | ||     |  ||| |  |||
Db       1856 ACGCCCTCAAGGACGCGGAGAAGTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACG 1915

Qy        305 ACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCA 344
              || || |||  ||| | | ||  |     |  | |||| |
Db       1916 CGCGCAAGGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955


RESULT 12
US-09-565-501A-101
; Sequence 101, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
```

```
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Peter Probst
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C6
;  CURRENT APPLICATION NUMBER: US/09/565,501A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  112
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-565-501A-101

  Query Match             4.3%;  Score 62.4;  DB 4;  Length 2735;
  Best Local Similarity   51.4%;  Pred. No. 2.1e-05;
  Matches  144;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         65 AAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCG 124
              |||||   |   |    |    |   ||||  | |    | |    | | |   ||  ||  ||||
Db       1676 AAGCCTACATCGATCCTGAGATCGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCA 1735

Qy        125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCG 184
                 || ||  | |||     ||||   |       ||||   ||||||| |    |   || |
Db       1736 AGGAGGATAAGTTCCCCGAGGCCGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTG 1795

Qy        185 CCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTG 244
               ||     |   |  |||   | | |||||    ||| ||||||||| ||     || ||  |
Db       1796 CCGAGCACACCTCCTACAGCAATCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACG 1855

Qy        245 AGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACT 304
               | ||  |      ||| ||  | ||     |||||||  | ||        |   ||| |  |||
Db       1856 ACGCCCTCAAGGACGCGGAGAAGTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACG 1915

Qy        305 ACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCA 344
                 || || |||  ||| | | ||  |      |   | ||| |
Db       1916 CGCGCAAGGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955


RESULT 13
US-09-639-206A-101
; Sequence 101, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C7
;  CURRENT APPLICATION NUMBER: US/09/639,206A
;  CURRENT FILING DATE:  2000-08-14
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-639-206A-101

  Query Match             4.3%;  Score 62.4;  DB 4;  Length 2735;
  Best Local Similarity   51.4%;  Pred. No. 2.1e-05;
  Matches  144;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         65 AAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCG 124
              |||||   |   |    |    |   ||||  | |    | |    | | |   ||  ||  ||||
Db       1676 AAGCCTACATCGATCCTGAGATCGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCA 1735

Qy        125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCG 184
                 || ||  | |||     ||||   |       ||||   ||||||| |    |   || |
Db       1736 AGGAGGATAAGTTCCCCGAGGCCGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTG 1795

Qy        185 CCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTG 244
               ||     |   |  |||   | | |||||    ||| ||||||||| ||     || ||  |
Db       1796 CCGAGCACACCTCCTACAGCAATCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACG 1855

Qy        245 AGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACT 304
               | ||  |      ||| ||  | ||     |||||||  | ||        |   ||| |  |||
Db       1856 ACGCCCTCAAGGACGCGGAGAAGTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACG 1915

Qy        305 ACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCA 344
                 || || |||  ||| | | ||  |      |   | ||| |
Db       1916 CGCGCAAGGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955


RESULT 14
US-09-874-923-101
; Sequence 101, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Brannon, Mark
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C8
;  CURRENT APPLICATION NUMBER: US/09/874,923
;  CURRENT FILING DATE:  2001-06-04
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-874-923-101

  Query Match             4.3%;  Score 62.4;  DB 4;  Length 2735;
  Best Local Similarity   51.4%;  Pred. No. 2.1e-05;
  Matches  144;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         65 AAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCG 124
              |||||   |   |    |    |   ||||  | |    | |    | | |   ||  ||  ||||
Db       1676 AAGCCTACATCGATCCTGAGATCGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCA 1735

Qy        125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCG 184
                 || ||  | |||     ||||   |       ||||   ||||||| |    |   || |
Db       1736 AGGAGGATAAGTTCCCCGAGGCCGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTG 1795

Qy        185 CCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTG 244
               ||     |   |  |||   | | |||||    ||| ||||||||| ||     || ||  |
Db       1796 CCGAGCACACCTCCTACAGCAATCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACG 1855

Qy        245 AGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACT 304
```

```
                | ||  |   ||| ||  | ||    ||||||||  | ||      |   ||| |  |||
Db       1856 ACGCCCTCAAGGACGCGGAGAAGTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACG 1915

Qy        305 ACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCA 344
              || || |||  ||| | | ||  |     |  | ||| |
Db       1916 CGCGCAAGGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955


RESULT 15
US-09-551-974A-94
; Sequence 94, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C5
;  CURRENT APPLICATION NUMBER: US/09/551,974A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 94
;   LENGTH: 3012
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-551-974A-94

  Query Match             4.3%;  Score 62.4;  DB 4;  Length 3012;
  Best Local Similarity   51.4%;  Pred. No. 2.2e-05;
  Matches  144;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy          65 AAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCG 124
               |||||   |  |   |   | ||||  | |   | |   | | |  ||  ||  ||||
Db        1676 AAGCCTACATCGATCCTGAGATCGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCA 1735

Qy         125 TCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCG 184
               || ||  | |||      ||||   |      ||||   |||||||| |   |   || |
Db        1736 AGGAGGATAAGTTCCCCGAGGCCGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTG 1795

Qy         185 CCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTG 244
               ||      |  |  |||   | | |||||    ||| ||||||||| ||   || ||  |
Db        1796 CCGAGCACACCTCCTACAGCAATCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACG 1855

Qy         245 AGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACT 304
               | ||  |   ||| ||  | ||    ||||||||  | ||      |   ||| |  |||
Db        1856 ACGCCCTCAAGGACGCGGAGAAGTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACG 1915

Qy         305 ACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCA 344
               || || |||  ||| | | ||  |     |  | ||| |
Db        1916 CGCGCAAGGGTCATGCTTACTTTTGGACCAAGCAGTACAA 1955


Search completed: April 23, 2005, 08:59:03
Job time : 261.537 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:35:27 ; Search time 6457.31 Seconds
(without alignments)
10918.214 Million cell updates/sec

Title: US-10-609-078-7
Perfect score: 1455
Sequence: 1 gaattcggcacgagcggcac.....aaaaaaaaaaaaaaaaaaaa 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

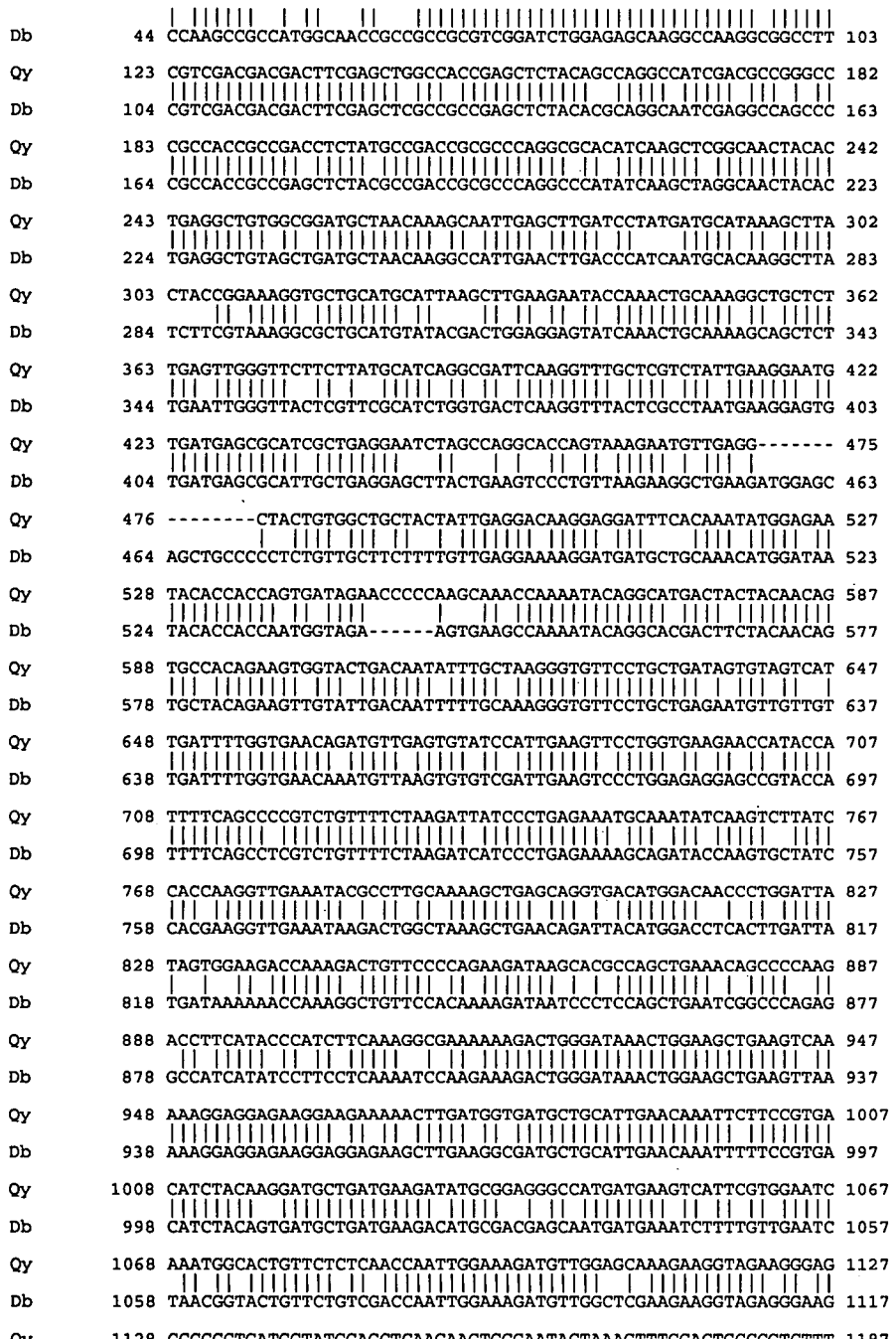
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	750.4	51.6	1409	8	AF439974	AF439974 Hordeum v
2	743.6	51.1	1222	8	AF192467	AF192467 Oryza sat
3	456.2	31.4	984	8	RID251317	AJ251317 Rubus ida
4	454.8	31.3	1113	8	AF494083	AF494083 Nicotiana
5	420	28.9	1366	8	AY085464	AY085464 Arabidops
6	419	28.8	1354	8	ATH318019	AJ318019 Arabidops
7	418.4	28.8	1108	8	AY150487	AY150487 Arabidops
8	418.4	28.8	1290	8	AF439976	AF439976 Arabidops
9	418.4	28.8	1400	8	AF428340	AF428340 Arabidops
10	418.4	28.8	1420	8	AF370229	AF370229 Arabidops
11	417.6	28.7	1077	6	AX412470	AX412470 Sequence
12	413.8	28.4	1197	8	AJ620883	AJ620883 Brassica
13	407.2	28.0	1068	8	AJ620882	AJ620882 Brassica
14	394.8	27.1	1291	8	AF439975	AF439975 Arabidops
15	393.2	27.0	1084	8	AY096395	AY096395 Arabidops
16	393.2	27.0	1348	8	AY063978	AY063978 Arabidops
17	285.8	19.6	968	8	ATH318020	AJ318020 Arabidops
18	261	17.9	582	8	AK059268	AK059268 Oryza sat
19	207.8	14.3	634	8	AF516180	AF516180 Nicotiana
20	207.8	14.3	634	8	AF516181	AF516181 Nicotiana
21	164.4	11.3	272	6	AR249214	AR249214 Sequence
22	163	11.2	1535	9	BC000911	BC000911 Homo sapi
23	163	11.2	1555	9	HSA344097	AJ344097 Homo sapi
24	163	11.2	1776	6	CQ490463	CQ490463 Sequence
25	163	11.2	1776	6	CQ496307	CQ496307 Sequence
26	160	11.0	1606	5	AJ720706	AJ720706 Gallus ga
27	159.6	11.0	1518	5	BC072118	BC072118 Xenopus l
28	159	10.9	1235	9	AY321358	AY321358 Homo sapi
29	156.4	10.7	1002	6	CQ724202	CQ724202 Sequence
30	156.4	10.7	1002	9	AF132856	AF132856 Homo sapi
31	156.4	10.7	1025	6	BD209970	BD209970 Human tra
32	156.4	10.7	2053	6	BD192630	BD192630 123 human
33	156.4	10.7	2053	6	AR184101	AR184101 Sequence
34	156.2	10.7	1002	9	BT009798	BT009798 Homo sapi
35	146.6	10.1	139201	8	AP002970	AP002970 Oryza sat
36	146.6	10.1	150022	8	AP003213	AP003213 Oryza sat
37	145.2	10.0	1314	5	CR761938	CR761938 Xenopus t
38	143.6	9.9	153865	8	AP002744	AP002744 Oryza sat
39	139.4	9.6	1104	10	BC009167	BC009167 Mus muscu
40	119	8.2	1089	9	AF068289	AF068289 Homo sapi
41	114.6	7.9	154558	8	AC107206	AC107206 Oryza sat
c 42	111	7.6	137843	9	AC138972	AC138972 Homo sapi
43	107.8	7.4	304	4	BTA344098	AJ344098 Bos tauru
44	92.2	6.3	1293	8	AK073324	AK073324 Oryza sat
45	90.8	6.2	1242	6	AX488766	AX488766 Sequence

ALIGNMENTS

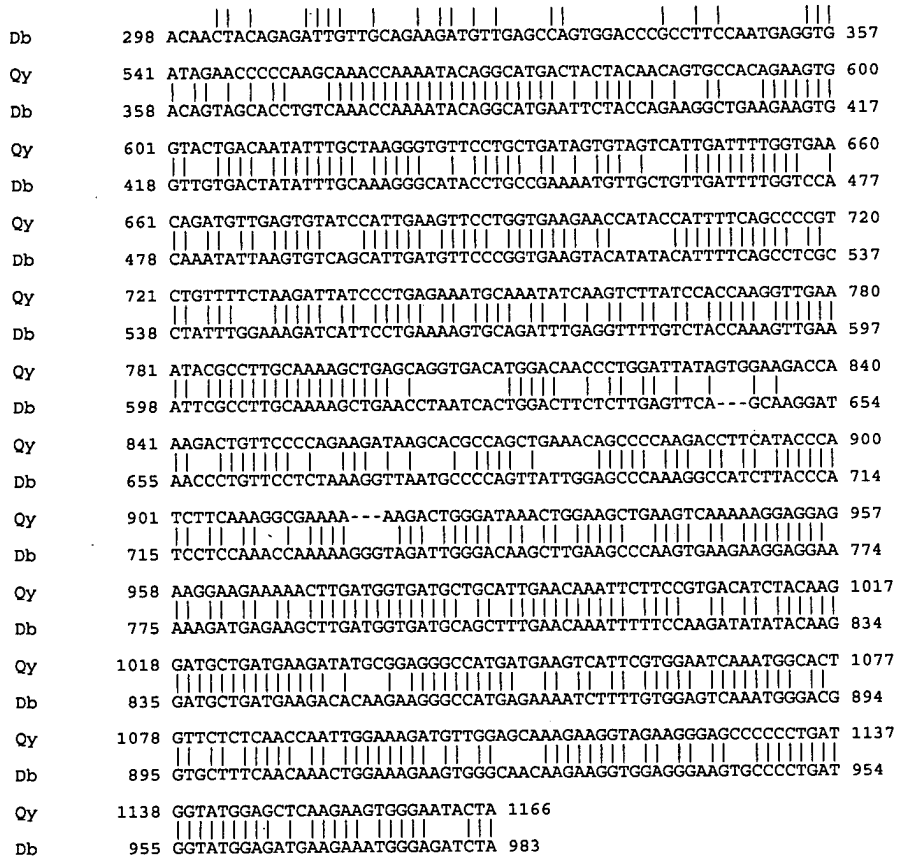
RESULT 1
AF439974
LOCUS AF439974 1409 bp mRNA linear PLN 06-NOV-2003
DEFINITION Hordeum vulgare SGT1 (SGT1) mRNA, complete cds.
ACCESSION AF439974
VERSION AF439974.1 GI:17017305
KEYWORDS .
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1409)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance
JOURNAL Science 295 (5562), 2073-2076 (2002)
MEDLINE 21893744
PUBMED 11847307
REFERENCE 2 (bases 1 to 1409)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich NR47UH, UK
FEATURES Location/Qualifiers
source 1. .1409
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
gene 1. .1409
/gene="SGT1"
CDS 9. .1130
/gene="SGT1"
/codon_start=1
/product="SGT1"
/protein_id="AAL33610.1"
/db_xref="GI:17017306"
/translation="MAAAAASDLESKAKEAFVDDDFELAAELYTQAIEAGPATAELYA
DRAQAHIKLGSYTEAVADANKAIELDPSMHKAYLRKGSACIKLEEYQTAKAALEVGSS
YASGDSRFTRLMKECDDRIAEEASQAPVKNAAAAVAPATSSGATTVVTEAEDQDGENM

```
                     ENAQPTVEVPSKPKYRHDYYNTPTEVVLTIFAKGVPADSVVVDFGEQMLSVSIELPGE
                     EPYHFQPRLFSKIVPDKCKYTVLSTKVEIRLAKAEPVTWTSLDYTGKPKAPQKINVPA
                     ESAQRPSYPSSKSKKDWDKLEAEVKKQEKDEKLDGDAALNKFFREIYSDADEDMRRAM
                     MKSFVESNGTVLSTNWKDVGKKTVEGSPPDGMELKKWEY"
ORIGIN

  Query Match           51.6%;  Score 750.4;  DB 8;  Length 1409;
  Best Local Similarity  79.5%;  Pred. No. 1.7e-120;
  Matches  927;  Conservative    0;  Mismatches  206;  Indels   33;  Gaps    2;

Qy         69 CGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGA 128
Db          5 CAGAATGGCCGCCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGA  64

Qy        129 CGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCAC 188
Db         65 CGACGACTTCGAGCTAGCCGCCGAGCTCTACACCCAGGCCATCGAGGCTGGCCCAGCCAC 124

Qy        189 CGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGC 248
Db        125 CGCGGAGCTCTACGCCGACCGAGCCCAGGCTCACATCAAGCTGGGCAGTTACACTGAGGC 184

Qy        249 TGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCG 308
Db        185 TGTAGCTGATGCCAACAAAGCAATTGAACTTGATCCCTCGATGCACAAAGCGTACCTTCG 244

Qy        309 GAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTT 368
Db        245 GAAGGGTTCTGCTTGCATCAAGCTGGAGGAATACCAAACTGCAAAGGCTGCTCTTGAAGT 304

Qy        369 GGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGA 428
Db        305 TGGTTCTTCCTATGCATCTGGTGACTCAAGGTTTACTCGTCTTATGAAGGAGTGTGATGA 364

Qy        429 GCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCT-- 486
Db        365 TCGTATTGCTGAGGAGGCTAGCCAGGCGCCAGTAAAGAATGCCGCTGCGGCTGTTGCTCC 424

Qy        487 ----------------------------GCTACTATTGAGGACAAGGAGGATTTCACAAA 518
Db        425 TGCTACATCTTCCGGGGCTACAACTGTGGTTACTGAAGCTGAGGACCAGGATGGTGAAAA 484

Qy        519 TATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTA 578
Db        485 TATGGAGAATGCACAGCCAACGGTAGAAGTGCCAAGCAAGCCCAAATACAGGCATGACTA 544

Qy        579 CTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAG 638
Db        545 CTACAATACTCCTACAGAAGTGGTACTGACTATATTTGCTAAGGGTGTTCCAGCTGACAG 604

Qy        639 TGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGA 698
Db        605 CGTGGTTGTTGACTTTGGTGAACAGATGCTGAGTGTCTCAATTGAACTTCCTGGTGAGGA 664

Qy        699 ACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCA 758
Db        665 ACCATACCATTTCCAGCCTCGTCTGTTTTCAAAGATCGTCCCAGATAAGTGCAAGTATAC 724

Qy        759 AGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAAC 818
Db        725 TGTATTGTCTACAAAGGTCGAAATACGCCTGGCAAAAGCTGAGCCAGTAACTTGGACATC 784

Qy        819 CCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAAC 878
Db        785 ACTGGATTATACTGGTAAACCAAAG---GCTCCTCAGAAGATAAATGTACCAGCTGAATC 841

Qy        879 AGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGC 938
Db        842 AGCCCAGAGGCCATCTTATCCTTCATCAAAATCCAAAAAGGACTGGGATAAGCTTGAGGC 901

Qy        939 TGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATT 998
Db        902 TGAAGTGAAAAAACAGGAGAAGGATGAGAAACTTGACGGTGATGCTGCATTGAACAAGTT 961

Qy        999 CTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATT 1058
Db        962 TTTCCGTGAAATTTACAGTGATGCTGATGAAGATATGCGTAGAGCAATGATGAAGTCTTT 1021

Qy       1059 CGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGT 1118
Db       1022 TGTGGAGTCTAATGGAACCGTTCTCTCAACCAACTGGAAAGATGTCGGGAAAAAGACGGT 1081

Qy       1119 AGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTG 1178
Db       1082 TGAAGGAAGCCCTCCTGATGGAATGGAGCTCAAGAAGTGGGAGTATTAATGAAACCAGAT 1141

Qy       1179 CCCGTCTTTTGTAAATCCAGGTCTTG 1204
Db       1142 CATCCATGTTGTAAATCCTGGTCGTG 1167


RESULT 2
AF192467
LOCUS       AF192467               1222 bp    mRNA    linear   PLN 16-DEC-1999
DEFINITION  Oryza sativa Sgt1 (Sgt1) mRNA, complete cds.
ACCESSION   AF192467
VERSION     AF192467.1  GI:6581057
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1222)
  AUTHORS   Shirasu,K., Lahaye,T., Tan,M.W., Zhou,F., Azevedo,C. and
            Schulze-Lefert,P.
  TITLE     A novel class of eukaryotic zinc-binding proteins is required for
            disease resistance signaling in barley and development in C.
            elegans
  JOURNAL   Cell 99 (4), 355-366 (1999)
  MEDLINE   20035956
   PUBMED   10571178
REFERENCE   2  (bases 1 to 1222)
  AUTHORS   Shirasu,K. and Schulze-Lefert,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-1999) The Sainsbury Laboratory, John Innes
            Center, Colney Lane, Norwich NR4 7UH, United Kingdom
FEATURES             Location/Qualifiers
     source          1. .1222
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:4530"
                     /note="corresponds to MAFF DNA Bank EST clone C62803"
     gene            1. .1222
                     /gene="Sgt1"
     CDS             54. .1157
                     /gene="Sgt1"
                     /note="similar to yeast Sgt1"
                     /codon_start=1
                     /product="Sgt1"
                     /protein_id="AAF18438.1"
                     /db_xref="GI:6581058"
                     /translation="MATAAASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYA
                     DRAQAHIKLGNYTEAVADANKAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYS
                     FASGDSRFTRLMKECDERIAEELTEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPP
                     MVEVKPKYRHDFYNSATEVVLTIFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQP
                     RLFSKIIPEKSRYQVLSTKVEIRLAKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRP
                     SYPSSKSKKDWDKLEAEVKKEEKEEKLEGDAALNKFFRDIYSDADEDMRRAMMKSFVE
                     SNGTVLSTNWKDVGSKKVEGSPPDGMELKKWEY"
ORIGIN

  Query Match           51.1%;  Score 743.6;  DB 8;  Length 1222;
  Best Local Similarity  79.5%;  Pred. No. 2.6e-119;
  Matches  913;  Conservative    0;  Mismatches  215;  Indels   21;  Gaps    2;

Qy         63 CAAAGCCGGAAGGGGCGCCATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTT 122
```



```
Db        44 CCAAGCCGCCATGGCAACCGCCGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTT 103

Qy       123 CGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCC 182

Db       104 CGTCGACGACGACTTCGAGCTCGCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCC 163

Qy       183 CGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACAC 242

Db       164 CGCCACCGCCGAGCTCTACGCCGACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACAC 223

Qy       243 TGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTA 302

Db       224 TGAGGCTGTAGCTGATGCTAACAAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTA 283

Qy       303 CTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCT 362

Db       284 TCTTCGTAAAGGCGCTGCATGTATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCT 343

Qy       363 TGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATG 422

Db       344 TGAATTGGGTTACTCGTTCGCATCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTG 403

Qy       423 TGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGG------- 475

Db       404 TGATGAGCGCATTGCTGAGGAGCTTACTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGC 463

Qy       476 --------CTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 527

Db       464 AGCTGCCCCCTCTGTTGCTTCTTTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAA 523

Qy       528 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 587

Db       524 TACACCACCAATGGTAGA------AGTGAAGCCAAAATACAGGCACGACTTCTACAACAG 577

Qy       588 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 647

Db       578 TGCTACAGAAGTTGTATTGACAATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGT 637

Qy       648 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 707

Db       638 TGATTTTGGTGAACAAATGTTAAGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCA 697

Qy       708 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 767

Db       698 TTTTCAGCCTCGTCTGTTTTCTAAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATC 757

Qy       768 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 827

Db       758 CACGAAGGTTGAAATAAGACTGGCTAAAGCTGAACAGATTACATGGACCTCACTTGATTA 817

Qy       828 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 887

Db       818 TGATAAAAAACCAAAGGCTGTTCCACAAAAGATAATCCCTCCAGCTGAATCGGCCCAGAG 877

Qy       888 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAA 947

Db       878 GCCATCATATCCTTCCTCAAAATCCAAGAAAGACTGGGATAAACTGGAAGCTGAAGTTAA 937

Qy       948 AAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGA 1007

Db       938 AAAGGAGGAGAAGGAGGAGAAGCTTGAAGGCGATGCTGCATTGAACAAATTTTTCCGTGA 997

Qy      1008 CATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATC 1067

Db       998 CATCTACAGTGATGCTGATGAAGACATGCGACGAGCAATGATGAAATCTTTTGTTGAATC 1057

Qy      1068 AAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAG 1127

Db      1058 TAACGGTACTGTTCTGTCGACCAATTGGAAAGATGTTGGCTCGAAGAAGGTAGAGGGAAG 1117

Qy      1128 CCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAAAGTTTGGACTGCCCGTCTTT 1187

Db      1118 CCCACCTGATGGGATGGAGCTTAAGAAATGGGAGTACTAAGATCTCGGNTCGTGCCATGC 1177

Qy      1188 TGTAAATCC 1196

Db      1178 TGTAAATTC 1186
```

```
RESULT 3
RID251317
LOCUS       RID251317               984 bp    mRNA    linear   PLN 24-NOV-1999
DEFINITION  Rubus idaeus partial mRNA for putative protein phosphatase.
ACCESSION   AJ251317
VERSION     AJ251317.1  GI:6468694
KEYWORDS    protein phosphatase.
SOURCE      Rubus idaeus
  ORGANISM  Rubus idaeus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rubus.
REFERENCE   1
  AUTHORS   Schroeder,G., Lurz,G. and Schroeder,J.
  TITLE     cDNA for protein phosphatase like protein from raspberry (Rubus
            idaeus)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 984)
  AUTHORS   Schroeder,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-1999) Schroeder G., Institut fuer Biologie II,
            Universitaet Freiburg, Schaenzlestr. 1, Freiburg D-79104, Germany
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Rubus idaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:32247"
     CDS             <1. .984
                     /codon_start=1
                     /product="putative protein phosphatase"
                     /protein_id="CAB61630.1"
                     /db_xref="GI:6468695"
                     /db_xref="UniProt/TrEMBL:Q9SM53"
                     /translation="PKNSELYSDRAQANMKSNNLTEAVADANKAIELDPSLFKAYLRK
                     GIACMKLEEYRTAKAALELGASLAPGESRFAKMIKECDELIAEENGEPPKQPMETTTT
                     EIVAEDVEPVDPPSNEVTVAPVKPKYRHEFYQKAEEVVVTIFAKGIPAENVAVDFGPQ
                     ILSVSIDVPGEVHIHFQPRLFGKIIPEKCRFEVLSTKVEIRLAKAEPNHWTSLEFSKD
                     NPVPLKVNAPVIGAQRPSYPSSKPKRVDWDKLEAQVKKEEKDEKLDGDAALNKFFQDI
                     YKDADEDTRRAMRKSFVESNGTVLSTNWKEVGNKKVEGSAPDGMEMKKWEI"
ORIGIN

  Query Match           31.4%;  Score 456.2;  DB 8;  Length 984;
  Best Local Similarity   68.5%;  Pred. No. 2.4e-69;
  Matches  677;  Conservative    0;  Mismatches  303;  Indels    9;  Gaps    3;

Qy       181 CCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTAC 240

Db         1 CCCAAGAACTCCGAGCTCTACTCCGACCGGGCTCAGGCCAACATGAAATCCAACAATCTC 60

Qy       241 ACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCT 300

Db        61 ACCGAGGCTGTTGCGGATGCGAACAAGGCCATTGAGTTGGACCCGTCTTTGTTCAAAGCC 120

Qy       301 TACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCT 360

Db       121 TACTTGCGCAAAGGAATTGCCTGCATGAAGCTCGAGGAATATAGAACTGCAAAGGCAGCC 180

Qy       361 CTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAA 420

Db       181 TTGGAACTTGGTGCTTCCTTGGCACCAGGGGAATCCAGATTCGCTAAAATGATAAAGGAG 240

Qy       421 TGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACT 480

Db       241 TGTGATGAGCTCATTGCAGAGGAAAATGGTGAGCCACC---GAAGCAACCAATGGAAACA 297

Qy       481 GTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTG 540
```

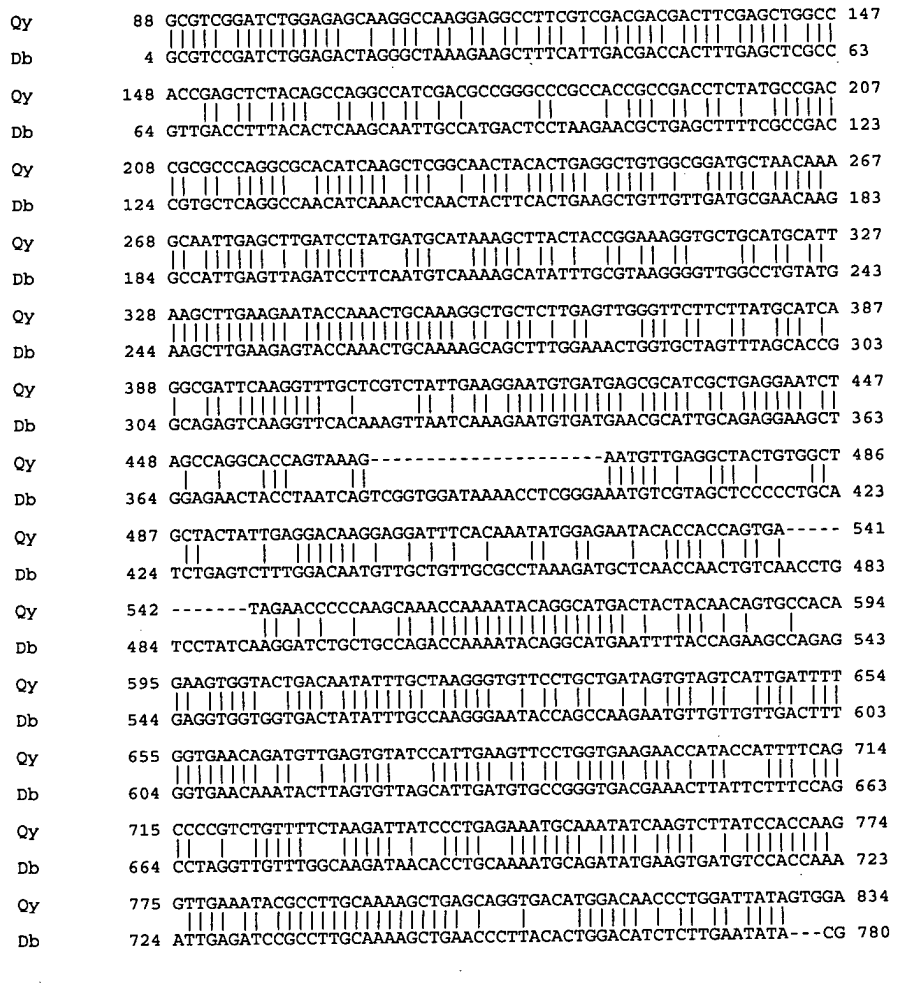


RESULT 4
AF494083
LOCUS AF494083 1113 bp mRNA linear PLN 30-OCT-2003
DEFINITION Nicotiana benthamiana SGT1 mRNA, complete cds.
ACCESSION AF494083
VERSION AF494083.1 GI:29468338
KEYWORDS .
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1113)
AUTHORS Liu,Y., Schiff,M., Serino,G., Deng,X.W. and Dinesh-Kumar,S.P.
TITLE Role of SCF ubiquitin-ligase and the COP9 signalosome in the N gene-mediated resistance response to Tobacco mosaic virus
JOURNAL Plant Cell 14 (7), 1483-1496 (2002)
MEDLINE 22115173
PUBMED 12119369
REFERENCE 2 (bases 1 to 1113)
AUTHORS Liu,Y., Schiff,M., Serino,G., Dinesh-Kumar,S.P. and Deng,X.-W.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) MCDB, Yale University, New Haven, CT 06520, USA
FEATURES Location/Qualifiers

source 1. .1113
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
CDS 1. .1113
/codon_start=1
/product="SGT1"
/protein_id="AAO85509.1"
/db_xref="GI:29468339"
/translation="MASDLETRAKEAFIDDHFELAVDLYTQAIAMTPKNAELFADRAQ
ANIKLNYFTEAVVDANKAIELDPSMSKAYLRKGLACMKLEEYQTAKAALETGASLAPA
ESRFTKLIKECDERIAEEAGELPNQSVDKTSGNVVAPPASESLDNVAVAPKDAQPTVN
LSYQGSAARPKYRHEFYQKPEEVVVTIFAKGIPAKNVVVDFGEQILSVSIDVPGDETY
SFQPRLFGKITPAKCRYEVMSTKIEIRLAKAEPLHWTSLEYTRASAVVQRPNVSSDAP
RPSYPSSKLRHVDWDKLEAEVKKEEKDEKLDGDAALNKFFRDIYKDADEDTRRAMMKS
FVESNGTVLSTNWKEVGTKKVEGSPPDGMELKKWEI"

ORIGIN

Query Match 31.3%; Score 454.8; DB 8; Length 1113;
Best Local Similarity 65.6%; Pred. No. 4.2e-69;
Matches 729; Conservative 0; Mismatches 347; Indels 36; Gaps 3;



```
Qy        835 AGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCA 894
Db        781 AGAGCGTCTGCTGTAGTGCAGAGGCCTAATGTGTCATCAGATGCCCCCCGGCCCAGTTAT 840

Qy        895 TACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAG 954
Db        841 CCTTCCTCGAAATTGAGACATGTGGATTGGGATAAATTAGAGGCTGAAGTGAAAAAGGAG 900

Qy        955 GAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTAC 1014
Db        901 GAAAAAGACGAAAAATTGGATGGAGATGCAGCATTGAACAAATTTTTCCGAGACATTTAC 960

Qy       1015 AAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGC 1074
Db        961 AAAGATGCCGATGAGGACACCAGAAGAGCCATGATGAAATCCTTTGTGGAATCTAATGGG 1020

Qy       1075 ACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCT 1134
Db       1021 ACTGTTTTGTCGACAAACTGGAAAGAAGTTGGCACAAAGAAGGTCGAAGGAAGCCCTCCG 1080

Qy       1135 GATGGTATGGAGCTCAAGAAGTGGGAATACTA 1166
Db       1081 GATGGCATGGAGCTGAAGAAATGGGAGATCTA 1112


RESULT 5
AY085464
LOCUS       AY085464               1366 bp    mRNA    linear   PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 152712 mRNA, complete sequence.
ACCESSION   AY085464
VERSION     AY085464.1  GI:21404174
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1366)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
  JOURNAL   Genome Biol. 3 (6), RESEARCH0029 (2002)
  MEDLINE   22088475
   PUBMED   12093376
REFERENCE   2  (bases 1 to 1366)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-Length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1366)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.

FEATURES             Location/Qualifiers
     source          1. .1366
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="152712"
     CDS             57. .1133
                     /codon_start=1
                     /product="SGT1a"
                     /protein_id="AAM62690.1"
                     /db_xref="GI:21553597"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match           28.9%;  Score 420;  DB 8;  Length 1366;
  Best Local Similarity 63.0%;  Pred. No. 4.7e-63;
  Matches  683;  Conservative    0;  Mismatches  395;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db         60 GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 119

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db        120 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 179

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        180 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 239

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        240 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 299

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        300 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 359

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        360 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 419

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        420 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        477 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 536

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        537 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 596

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        597 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 656

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 716

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        717 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 776

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
```

```
Db        777 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 836

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927

Db        837 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987

Db        894 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 953

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047

Db        954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107

Db       1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167

Db       1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133

Qy       1168 AGTT 1171

Db       1134 TCTT 1137


RESULT 6
ATH318019
LOCUS       ATH318019              1354 bp    mRNA    linear   PLN 23-OCT-2003
DEFINITION  Arabidopsis thaliana mRNA for SGT1-like protein (sgt1a gene).
ACCESSION   AJ318019
VERSION     AJ318019.1  GI:30524963
KEYWORDS    SGT1-like protein; sgt1a gene.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Pawlowski,J., Holzmann,M., Berney,C., Fahrni,J., Gooday,A.J.,
            Cedhagen,T., Habura,A. and Bowser,S.S.
  TITLE     The evolution of early Foraminifera
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11494-11498 (2003)
  MEDLINE   22882917
   PUBMED   14504394
REFERENCE   2  (bases 1 to 1354)
  AUTHORS   Shen,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) Shen W.H., Biologie Cellulaire, IBMP/CNRS,
            12, rue du General Zimmer, 67084 Strasbourg, FRANCE
FEATURES             Location/Qualifiers
     source          1. .1354
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1. .1354
                     /gene="sgt1a"
     CDS             40. .1116
                     /gene="sgt1a"
                     /function="proteolysis"
                     /codon_start=1
                     /product="SGT1-like protein"
                     /protein_id="CAC85266.1"
                     /db_xref="GI:30524964"
                     /db_xref="UniProt/TrEMBL:Q9SUT5"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match           28.8%;  Score 419;  DB 8;  Length 1354;
  Best Local Similarity 62.9%;  Pred. No. 7e-63;
  Matches  682;  Conservative    0;  Mismatches  396;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147

Db         43 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 102

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207

Db        103 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 162

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267

Db        163 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 222

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327

Db        223 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 282

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387

Db        283 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 342

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447

Db        343 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 402

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507

Db        403 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 459

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567

Db        460 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 519

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627

Db        520 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 579

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687

Db        580 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 639

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747

Db        640 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 699

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807

Db        700 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 759

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867

Db        760 ACNTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 819

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927

Db        820 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 876

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987

Db        877 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 936

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047

Db        937 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 996

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
```

```
Db        997 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1056

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167

Db       1057 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1116

Qy       1168 AGTT 1171

Db       1117 TCTT 1120
```

RESULT 7
AY150487
LOCUS AY150487 1108 bp mRNA linear PLN 23-SEP-2002
DEFINITION Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete cds.
ACCESSION AY150487
VERSION AY150487.1 GI:23297701
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1108)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1108)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES Location/Qualifiers
source 1. .1108
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U09567"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
gene 1. .1108
/gene="At4g11260"
CDS 1. .1077
/gene="At4g11260"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AAN12904.1"
/db_xref="GI:23297702"
/translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
WKEVGTKKVESTPPDGMELKKWEY"
3'UTR 1078. .1108
/gene="At4g11260"
ORIGIN

Query Match 28.8%; Score 418.4; DB 8; Length 1108;
Best Local Similarity 62.9%; Pred. No. 8.9e-63;
Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;

```
Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147

Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207

Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267

Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327

Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387

Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447

Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507

Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567

Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627

Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687

Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747

Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807

Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720
```

```
Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077

Qy       1168 AGTT 1171
Db       1078 TCTT 1081
```

RESULT 8
AF439976
LOCUS AF439976 1290 bp mRNA linear PLN 06-NOV-2003
DEFINITION Arabidopsis thaliana SGT1b (SGT1b) mRNA, complete cds.
ACCESSION AF439976
VERSION AF439976.1 GI:17017309
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1290)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance
JOURNAL Science 295 (5562), 2073-2076 (2002)
MEDLINE 21893744
PUBMED 11847307
REFERENCE 2 (bases 1 to 1290)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich NR47UH, UK
FEATURES Location/Qualifiers
source 1. .1290
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
gene 1. .1290
/gene="SGT1b"
CDS 1. .1077
/gene="SGT1b"
/codon_start=1
/product="SGT1b"
/protein_id="AAL33612.1"
/db_xref="GI:17017310"
/translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK

```
                                                DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                                                WKEVGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match             28.8%;  Score 418.4;  DB 8;  Length 1290;
  Best Local Similarity   62.9%;  Pred. No. 8.9e-63;
  Matches  682;  Conservative    0;  Mismatches  396;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
```

```
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167

Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077

Qy       1168 AGTT 1171

Db       1078 TCTT 1081


RESULT 9
AF428340
LOCUS       AF428340               1400 bp    mRNA    linear   PLN 18-OCT-2001
DEFINITION  Arabidopsis thaliana AT4g11260/F8L21_50 mRNA, complete cds.
ACCESSION   AF428340
VERSION     AF428340.1  GI:16226817
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1400)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1400)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
            Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J. Bowser,L.,
            Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S.,
            Davis,R.W.,  Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1400
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL09-50-H05(R14746)"
                     /ecotype="Columbia"
     5'UTR           1. .131
     CDS             132. .1208
                     /note="unknown protein"
                     /codon_start=1
                     /product="AT4g11260/F8L21_50"
                     /protein_id="AAL16270.1"
                     /db_xref="GI:16226818"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
     misc_difference 703
                     /note="compared to genome"
                     /replace="g"
     3'UTR           1209. .1400
ORIGIN

  Query Match             28.8%;  Score 418.4;  DB 8;  Length 1400;
  Best Local Similarity   62.9%;  Pred. No. 8.9e-63;
  Matches  682;  Conservative    0;  Mismatches  396;  Indels    6;  Gaps    2;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147

Db        135 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 194

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207

Db        195 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 254

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267

Db        255 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 314

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327

Db        315 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 374

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387

Db        375 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 434

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447

Db        435 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 494

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507

Db        495 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 551

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567

Db        552 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 611

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627

Db        612 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 671

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687

Db        672 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 731

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747

Db        732 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 791

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807

Db        792 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 851
```

```
Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
              || ||| |  |||| || ||| ||  | | | ||    | | ||| |    |  | ||
Db        852 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 911

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
              ||||        | |  |||||    |||||||||||| ||| |   ||||||||||||
Db        912 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 968

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
              ||  ||||||||||||| || ||  |||||||||| || || |||||||| ||||| ||
Db        969 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 1028

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
              ||||||||||| ||| | ||||| |||  |  ||| |||||||| ||| || |||| |||
Db       1029 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1088

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
              |  || ||||| |  || || ||||| || || || || || || |||||||| |||||
Db       1089 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1148

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              | ||||| || ||  | |  || || ||||| ||||||||||||||||||| || | |
Db       1149 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1208

Qy       1168 AGTT 1171
                ||
Db       1209 TCTT 1212
```

RESULT 10
AF370229

LOCUS AF370229 1420 bp mRNA linear PLN 18-SEP-2002

DEFINITION Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete cds.

ACCESSION AF370229

VERSION AF370229.1 GI:13877932

KEYWORDS FLI_CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1420)

AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE Arabidopsis Full Length cDNA Clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1420)

AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE Direct Submission

JOURNAL Submitted (18-APR-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES Location/Qualifiers

```
     source            1. .1420
                       /organism="Arabidopsis thaliana"
                       /mol_type="mRNA"
                       /db_xref="taxon:3702"
                       /chromosome="4"
                       /clone="RAFL04-14-O07 (R09567)"
                       /ecotype="Columbia"
                       /note="This clone is in a modified pBluescript vector
                       (Lambda ZAP) as a XhoI/SstI insert."
     gene              1. .1420
                       /gene="At4g11260"
     5'UTR             1. .136
                       /gene="At4g11260"
     misc_difference   32. .34
                       /gene="At4g11260"
                       /note="not present in genomic sequence"
     CDS               137. .1213
                       /gene="At4g11260"
                       /codon_start=1
                       /evidence=experimental
                       /product="unknown protein"
                       /protein_id="AAK44044.1"
                       /db_xref="GI:13877933"
                       /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                       ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                       EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                       RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                       EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                       DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                       WKEVGTKKVESTPPDGMELKKWEY"
     3'UTR             1214. .1420
                       /gene="At4g11260"
```

ORIGIN

Query Match 28.8%; Score 418.4; DB 8; Length 1420;
Best Local Similarity 62.9%; Pred. No. 8.9e-63;
Matches 682; Conservative 0; Mismatches 396; Indels 6; Gaps 2;

```
Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              ||   |||  | |      || || || || || ||  | || || ||||||||  | ||
Db        140 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 199

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
                ||  | |||  | | ||||| |||    | |||      ||||| | |||  |||||
Db        200 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 259

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              || || |||||  |||||||  |||  |||| ||||| ||||| |  |||||  ||||||
Db        260 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 319

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              || |||||| | || || | |  |||||  ||   | | || ||  |||| || ||
Db        320 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 379

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
              ||||| ||||||||     ||||| || || || || ||   |||  |||||  |||| |
Db        380 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 439

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
              ||  ||| ||||         |  | | ||||| |||   || || || || |||
Db        440 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 499
```

```
Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        500 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 556

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        557 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 616

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        617 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 676

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        677 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 736

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        737 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 796

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        797 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 856

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        857 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 916

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        917 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 973

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        974 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 1033

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db       1034 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1093

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db       1094 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1153

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
Db       1154 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1213

Qy       1168 AGTT 1171
Db       1214 TCTT 1217
```

RESULT 11
AX412470
LOCUS AX412470 1077 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 234 from Patent WO0222675.
ACCESSION AX412470
VERSION AX412470.1 GI:21444928
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
TITLE Plant genes, the expression of which are altered by pathogen infection
JOURNAL Patent: WO 0222675-A 234 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES Location/Qualifiers
source 1. .1077
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN

Query Match 28.7%; Score 417.6; DB 6; Length 1077;
Best Local Similarity 63.0%; Pred. No. 1.2e-62;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

```
Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
```

```
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107

Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167

Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077
```

RESULT 12
AJ620883
LOCUS AJ620883 1197 bp mRNA linear PLN 16-JAN-2004
DEFINITION Brassica oleracea mRNA for SGT1-like protein (sgt1b-2 gene).
ACCESSION AJ620883
VERSION AJ620883.1 GI:40974916
KEYWORDS SGT1-like protein; sgt1b-2 gene.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Yemm,A.A., Holub,E. and Tor,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1197)
AUTHORS Yemm,A.A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research, Horticultural Research International, Wellesbourne, Warwick, UNITED KINGDOM

FEATURES Location/Qualifiers

```
     source          1. .1197
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /variety="Senna"
                     /db_xref="taxon:3712"
                     /tissue_type="leaf"
     gene            1. .1197
                     /gene="sgt1b-2"
     CDS             1. .1065
                     /gene="sgt1b-2"
                     /function="role in disease resistance"
                     /codon_start=1
                     /evidence=experimental
                     /product="SGT1-like protein"
                     /protein_id="CAF06581.1"
                     /db_xref="GI:40974917"
                     /translation="MASELAEKAKEAFLDDDFDVAADLYSKAIDLDPSCASFFADRAQ
                     ANIKLLNFTEAVADANKAIELEPTLAKAYLRKGAACMKLEEYATAKAALEKGASVAPN
                     ESKFEKMIDECNLLIAEEEKDLVQQVPPTLPSSSTTPLAIAADAPPAAPAKPMFRHEF
                     YQKPEEVGVAIFAKGIPKQNVNVEFGDQILSVVIDVAGEEAYHFQPRLFGKIIPEKCR
                     YEVLSTKVEIRLAKAEIVTWASLEYGKGQALLPKPNVASAVSQRPVYPSSKPGKDWDK
                     LEAEVKKQEKDEKLDGDAAMNKFFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWKEV
                     GTKKVESTPPDGMELKKWEY"
```

ORIGIN

Query Match 28.4%; Score 413.8; DB 8; Length 1197;
Best Local Similarity 63.6%; Pred. No. 5.6e-62;
Matches 687; Conservative 1; Mismatches 373; Indels 20; Gaps 3;

```
Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147

Db          4 GCGAGCGAACTAGCGGAAAAAGCTAAAGAGGCGTTCTTAGATGACGACTTCGATGTGGCT 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207

Db         64 GCTGACTTATACTCCAAAGCCATTGACTTGGATCCCTCTTGCGCCTCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267

Db        124 CGTGCTCAAGCCAACATCAAACTCCTTAACTTCACCGAAGCTGTTGCAGATGCCAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327

Db        184 GCCATCGAGTTGGAGCCTACTTTGGCCAAAGCTTATCTCAGAAAGGGCGCTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387

Db        244 AAGCTAGAAGAGTATGCTACTGCTAAAGCAGCTCTTGAGAAGGGCGCTTCCGTTGCACCA 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447

Db        304 AATGAATCCAAGTTTGAGAAGATGATAGATGAATGCAACCTTCTCATTGCAGAAGAA--- 360

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507

Db        361 ------------GAGAAAGATTTGGTTCAACAGGTGCCACCGACTTTGCCTTCAAGCTC 407

Qy        508 GATTTCACAAATATGGAGAATA-CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATA 566

Db        408 TACAACACCACTAGCAATCGCAGCTGATGCTCCTCCTGCAGCACCTGCCAAACCCATGTT 467

Qy        567 CAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGT 626

Db        468 CAGACACGAGTTCTACCAGAAGCCGGAAGAAGTGGGGGTGGCAATTTTCGCTAAAGGGAT 527

Qy        627 TCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGT 686

Db        528 CCCAAAGCAGAACGTGAATGTAGAGTTTGGTGACCAGATTCTGAGTGTTGTGATTGATGT 587

Qy        687 TCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAA 746

Db        588 TGCTGGAGAAGAAGCTTATCATTTCCAGCCAAGATTGTTCGGGAAGATAATACCAGAGAA 647

Qy        747 ATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGT 806

Db        648 GTGCAGATATGAGGTATTGTCGACCAAAGTTGAGATCCGTCTCGCAAAAGCAGAGATAGT 707

Qy        807 GACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCAC 866

Db        708 CACTTGGGCCTCCCTTGAATA---TGGCAAAGGGCAAGCTCTTCTGCCTAAACCCAATGT 764

Qy        867 GCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGA 926

Db        765 TGCATCAGCAGTGTCTCAGAGGCCGGTGTACCCATCCTCTAAGCCGGGAAAAGACTGGGA 824

Qy        927 TAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGC 986

Db        825 CAAGCTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTAGATGGAGATGCAGC 884

Qy        987 ATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCAT 1046

Db        885 TATGAACAAGTTTTTCAGCGACATATACCAGAGTGCAGATGAGGACATGAGGCGAGCCAT 944

Qy       1047 GATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGG 1106

Db        945 GAACAAATCATTTGCAGAGTCCAATGGGACGGTGCTGTCTACAAACTGGAAAGAGGTTGG 1004

Qy       1107 AGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTA 1166

Db       1005 GACAAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTSAAGAAGTGGGAGTATTG 1064

Qy       1167 A 1167

Db       1065 A 1065
```

RESULT 13
AJ620882
LOCUS AJ620882 1068 bp mRNA linear PLN 16-JAN-2004
DEFINITION Brassica oleracea mRNA for SGT1-like protein (sgt1b-1 gene).
ACCESSION AJ620882
VERSION AJ620882.1 GI:40974914

```
KEYWORDS    SGT1-like protein; sgt1b-1 gene.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Yemm,A.A., Holub,E. and Tor,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1068)
  AUTHORS   Yemm,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,
            Horticultural Research International, Wellesbourne, Warwick, UNITED
            KINGDOM
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /variety="Senna"
                     /db_xref="taxon:3712"
                     /tissue_type="leaf"
     gene            1. .1068
                     /gene="sgt1b-1"
     CDS             1. .1068
                     /gene="sgt1b-1"
                     /function="role in disease resistance"
                     /codon_start=1
                     /evidence=experimental
                     /product="SGT1-like protein"
                     /protein_id="CAF06580.1"
                     /db_xref="GI:40974915"
                     /translation="MASELAEKAKEAFLEDDFDVAVDLYSRAIDLDPNCAAFFADRAQ
                     ANIKILNFTEAVADANKAIELEPTLSKAYLRKGTACMKLEEYSTAKAALQKGASVAPN
                     ESKFNKLIDECNLHIAEEEKDLAQQMPPTLPSSSTTPPLATAADSPPAAPAKPMFRHE
                     FYQKPEEVVVTVFAKGIPKQNLNVEFGDQILSVVIDVAGEEAYHFQPRLFGKIIPDKC
                     RYEVLSTKVEIRLAKAEIITWASLEYVKGQALLPKPNVASAVSQRPVYPSSKPAKDWD
                     KLEAEVKKQEKDEKLDGDAAMNKFFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWKE
                     VGTKKVESTPPDGMELKKWEY"
ORIGIN
```

Query Match 28.0%; Score 407.2; DB 8; Length 1068;
Best Local Similarity 63.1%; Pred. No. 7.9e-61;
Matches 682; Conservative 0; Mismatches 383; Indels 15; Gaps 3;

```
Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
Db          4 GCGAGCGAACTAGCGGAAAAAGCTAAAGAGGCTTTCTTAGAGGACGACTTCGATGTCGCC 63

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
Db         64 GTTGACTTATACTCCAGAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
Db        124 CGTGCTCAGGCCAACATCAAAATCCTTAACTTTACCGAAGCTGTTGCAGATGCCAACAAA 183

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
Db        184 GCCATTGAGTTGGAGCCTACGTTGTCTAAAGCTTACCTCAGAAAAGGCACTGCTTGTATG 243

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
Db        244 AAGCTAGAGGAATACAGTACTGCTAAAGCAGCCCTTCAAAAGGGCGCTTCCGTTGCACCA 303

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
Db        304 AATGAATCAAAGTTTAACAAGTTGATAGATGAATGCAATCTTCACATTGCAGAAGAAGAG 363

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
Db        364 AAAGATTTGGCTCAACAGA------TGCCGCCCACCTTGCCTTCAAGCTCTACCACTCCA 417

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
Db        418 CCACTAGCAACTGCAGCTGATTCTCCTCCTG------CAGCACCTGCCAAACCCATGTTC 471

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
Db        472 AGACACGAGTTCTACCAGAAGCCGGAAGAAGTGGTGGTGACAGTTTTCGCTAAAGGGATA 531

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
Db        532 CCAAAGCAGAACTTGAATGTCGAGTTTGGTGATCAGATTCTGAGTGTTGTGATTGATGTT 591

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
Db        592 GCTGGAGAGGAAGCTTATCACTTCCAGCCAAGATTGTTTGGGAAGATAATACCAGACAAG 651

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
Db        652 TGCAGATATGAAGTATTGTCAACCAAAGTTGAGATCCGTCTCGCAAAAGCAGAGATAATC 711

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
Db        712 ACTTGGGCCTCCCTTGAATA---TGTCAAAGGGCAAGCTCTTCTGCCTAAACCCAATGTC 768

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
Db        769 GCATCAGCGGTTTCGCAGAGGCCGGTGTACCCATCATCTAAGCCGGCGAAAGACTGGGAC 828

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
Db        829 AAGCTGGAGGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTAGATGGAGATGCAGCT 888

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
Db        889 ATGAACAAGTTTTTCAGCGATATATACCAGAGTGCTGATGAGGATATGAGGCGAGCCATG 948

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
Db        949 AACAAATCATTTGCTGAGTCAAATGGGACAGTGCTGTCTACGAACTGGAAAGAGGTTGGG 1008

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
Db       1009 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGTATGGAGCTCAAGAAGTGGGAGTATTGA 1068
```

```
RESULT 14
AF439975
LOCUS       AF439975               1291 bp    mRNA    linear   PLN 06-NOV-2003
DEFINITION  Arabidopsis thaliana SGT1a (SGT1a) mRNA, complete cds.
ACCESSION   AF439975
VERSION     AF439975.1  GI:17017307
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1291)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     The RAR1 interactor SGT1, an essential component of R
            gene-triggered disease resistance
  JOURNAL   Science 295 (5562), 2073-2076 (2002)
  MEDLINE   21893744
   PUBMED   11847307
REFERENCE   2  (bases 1 to 1291)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich NR47UH, UK
FEATURES             Location/Qualifiers
     source          1. .1291
```

```
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1. .1291
                     /gene="SGT1a"
     CDS             77. .1129
                     /gene="SGT1a"
                     /codon_start=1
                     /product="SGT1a"
                     /protein_id="AAL33611.1"
                     /db_xref="GI:17017308"
                     /translation="MAKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQ
                     AYIKLESFTEAVADANKAIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPS
                     ESKFKKLIDECNFLITEEEKDLVQPVPSTLPSSVTAPPVSELDVTPTAKYRHEYYQKP
                     EEVVVTVFAKGIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVL
                     STKIEIRLAKADIITWASLEHGKGPAVLPKPNVSSEVSQRPAYPSSKKVKDWDKLEAE
                     VKKQEKDEKLEGDAALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKT
                     IESTPPDGMELKKWEI"
ORIGIN

  Query Match           27.1%;  Score 394.8;  DB 8;  Length 1291;
  Best Local Similarity   62.8%;  Pred. No. 1.1e-58;
  Matches  678;  Conservative    0;  Mismatches  372;  Indels   30;  Gaps    3;

Qy         88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
              |||  ||| || |     ||||| || ||||| ||||| ||||| ||||||||  | ||
Db         80 GCGAAGGAGCTTGCTGATAAGGCTAAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT 139

Qy        148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
                ||  | |||  | | ||||| |||   |  ||       ||| ||  |||  || ||
Db        140 GTTGACTTGTACTCCAAAGCCATTGACTTGGATCCTAATTGCGCTGAGTTCTTCGCTGAT 199

Qy        208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
              || || |||||  | ||||| |||| | || |||||||| ||||| ||||| ||||||
Db        200 CGTGCTCAGGCCTATATCAAACTCGAAAGCTTCACTGAGGCCGTGGCAGATGCGAACAAA 259

Qy        268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
              |||||||| | ||||||    ||   |||||||||    | |||||  |||| || ||
Db        260 GCAATTGAGTTGGATCCTTCATTGACCAAAGCTTACCTAAGAAAAGGAACTGCCTGTATG 319

Qy        328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
              |||||||||| || |  |||| ||  | ||||||||   |||| | |||    | |
Db        320 AAGCTTGAAGAGTATCGGACTGCTAAAACAGCTCTTGAAAAGGGTGCCTCTATCACGCCG 379

Qy        388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
              | || || |  |||        |  | | |||||  ||    |  ||| | || |||
Db        380 AGTGAATCCAAATTTAAGAAGTTGATAGATGAATGCAATTTTCTAATCACAGAAGAA--- 436

Qy        448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
                  | | |   |||   | || |  ||       |||    ||||
Db        437 ------------GAGAAAGATTTGGTTCAACCGGTGCCTTCGACTTTGCCTTCAAGTGT 483

Qy        508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
              ||   ||| |           ||  | | |   | ||   | |    | | | || |||
Db        484 GACAGCACCA-----------CCAGTATCTGAACTTGATGTCACCCCTACAGCCAAGTAC 532

Qy        568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
              |||||  || |||||| | |     |||| ||  | ||  | ||||| || ||  |
Db        533 AGGCACGAGTACTACCAAAAGCCAGAGGAAGTTGTGGTAACTGTTTTTGCAAAAGGAATA 592

Qy        628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
              ||     | | |||   ||| || |||||||||| ||  |||||||    |||| |||
Db        593 CCCAAGCAGAATGTTAACATCGACTTTGGTGAACAAATTCTGAGTGTTGTGATTGAGGTT 652

Qy        688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
              ||||| || ||  | ||  || | || ||   |  |||||   |||| || || || ||
Db        653 CCTGGAGAGGATGCGTATTATCTCCAACCGAGATTGTTTGGAAAGATAATACCAGATAAG 712

Qy        748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
              |||||||| |||| || || ||||  |||| || || ||||||||||  ||     |
Db        713 TGCAAATATGAAGTATTGTCGACCAAAATTGAGATCCGTCTTGCAAAAGCCGATATAATC 772

Qy        808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
              |||||| |  |||| ||  |  |   | | |||  |  | | ||  ||   |    | |
Db        773 ACATGGGCCTCCCTCGAACACGGCAAAGGGCCAGCGGTTTTGCCAAAGCCTAATGTCTCA 832

Qy        868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
              |||  |   ||   |  || ||  | || |||||||| |||      || ||||||||
Db        833 TCAGAGGTTTCA---CAGAGGCCAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGGAC 889

Qy        928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
              || |||||||||||||| || ||| |||| || || || || || || || || || ||
Db        890 AAGCTGGAAGCTGAAGTGAAGAAGCAGGAAAAAGATGAGAAGCTGGAAGGAGACGCAGCT 949

Qy        988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
              ||||||||||||||||| || || ||| || |||| ||||| |||||| || | || |||
Db        950 TTGAACAAATTCTTCCGGGAGATATACCAGAATGCGGATGAGGATATGAGGCGAGCGATG 1009

Qy       1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
              |  || ||||| |||||||| ||||| || || || ||||| || ||| |||| |||||
Db       1010 AGCAAATCATTTGTGGAATCGAATGGGACAGTGCTGTCAACAAACTGGCAAGAGGTTGGG 1069

Qy       1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
              | ||||   | ||  | |  || || ||||| |||||||||||||| |||||   || |
Db       1070 ACTAAGACAATCGAGAGTACTCCTCCGGATGGCATGGAGCTCAAGAAATGGGAGATCTGA 1129


RESULT 15
AY096395
LOCUS       AY096395               1084 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana putative phosphatase (At4g23570) mRNA,
            complete cds.
ACCESSION   AY096395
VERSION     AY096395.1  GI:20465860
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1084)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1084)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
```

Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

```
FEATURES             Location/Qualifiers
     source          1. .1084
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="U11621"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1. .1084
                     /gene="At4g23570"
     CDS             1. .1053
                     /gene="At4g23570"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative phosphatase"
                     /protein_id="AAM20035.1"
                     /db_xref="GI:20465861"
                     /translation="MAKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQ
                     AYIKLESFTEAVADANKAIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPS
                     ESKFKKLIDECNFLITEEEKDLVQPVPSTLPSSVTAPPVSELDVTPTAKYRHEYYQKP
                     EEVVVTVFAKGIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVL
                     STKIEICLAKADIITWASLEHGKGPAVLPKPNVSSEVSQRPAYPSSKKVKDWDKLEAE
                     VKKQEKDEKLEGDAALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKT
                     IESTPPDGMELKKWEI"
     3'UTR           1054. .1084
                     /gene="At4g23570"
ORIGIN
```

Query Match 27.0%; Score 393.2; DB 8; Length 1084;
Best Local Similarity 62.7%; Pred. No. 2.1e-58;
Matches 677; Conservative 0; Mismatches 373; Indels 30; Gaps 3;

```
Qy        88 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 147
             |||  ||| || |     ||||| || ||||| ||||| ||||| ||||||||  | ||
Db         4 GCGAAGGAGCTTGCTGATAAGGCTAAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT 63

Qy       148 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 207
               ||  | |||  | | ||||| |||   |  ||      ||| ||  |||  || ||
Db        64 GTTGACTTGTACTCCAAAGCCATTGACTTGGATCCTAATTGCGCTGAGTTCTTCGCTGAT 123

Qy       208 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 267
             || || |||||  | ||||| ||||  | || |||||||| ||||| ||||| ||||||
Db       124 CGTGCTCAGGCCTATATCAAACTCGAAAGCTTCACTGAGGCCGTGGCAGATGCGAACAAA 183

Qy       268 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 327
             |||||||||  | ||||||     ||   |||||||||     | |||||  |||| || ||
Db       184 GCAATTGAGTTGGATCCTTCATTGACCAAAGCTTACCTAAGAAAAGGAACTGCCTGTATG 243

Qy       328 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 387
             ||||||||||| || |  ||||| ||  | ||||||||   |||| | |||     |  |
Db       244 AAGCTTGAAGAGTATCGGACTGCTAAAACAGCTCTTGAAAAGGGTGCCTCTATCACGCCG 303

Qy       388 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 447
              | || || |  |||       |  |  | |||||  ||   |  ||| | || |||
Db       304 AGTGAATCCAAATTTAAGAAGTTGATAGATGAATGCAATTTTCTAATCACAGAAGAA--- 360

Qy       448 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 507
                       | | |   |||   | || |  ||         |||    ||||
Db       361 ------------GAGAAAGATTTGGTTCAACCGGTGCCTTCGACTTTGCCTTCAAGTGT 407

Qy       508 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 567
             ||   ||| |           ||  | | |   | ||   | |    | | | || |||
Db       408 GACAGCACCA-----------CCAGTATCTGAACTTGATGTCACCCCTACAGCCAAGTAC 456

Qy       568 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 627
             |||||  || |||||| | |    |     |||||  ||   | ||  | |||||  || ||  |
Db       457 AGGCACGAGTACTACCAAAAGCCAGAGGAAGTTGTGGTAACTGTTTTTGCAAAAGGAATA 516

Qy       628 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 687
             ||      | | |||    ||| || ||||||||||| ||  |||||||     ||||| |||
Db       517 CCCAAGCAGAATGTTAACATCGACTTTGGTGAACAAATTCTGAGTGTTGTGATTGAGGTT 576

Qy       688 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 747
             |||||  || ||   | ||  || | || ||  |  |||||    ||||| || || || ||
Db       577 CCTGGAGAGGATGCGTATTATCTCCAACCGAGATTGTTTGGAAAGATAATACCAGATAAG 636

Qy       748 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 807
             |||||||||  |||| || || |||||  |||| ||  | ||||||||||| ||      |
Db       637 TGCAAATATGAAGTATTGTCGACCAAAATTGAGATCTGTCTTGCAAAAGCCGATATAATC 696

Qy       808 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 867
             ||||||  |  |||| ||  |  |   | | ||| |  | | ||  ||   |    | |
Db       697 ACATGGGCCTCCCTCGAACACGGCAAAGGGCCAGCGGTTTTGCCAAAGCCTAATGTCTCA 756

Qy       868 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 927
             |||  |   ||  |  || ||  | || ||||||||  |||      || ||||||||
Db       757 TCAGAGGTTTCA---CAGAGGCCAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGGAC 813

Qy       928 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 987
             || |||||||||||||| || ||| |||| || || || || || || || || || ||
Db       814 AAGCTGGAAGCTGAAGTGAAGAAGCAGGAAAAAGATGAGAAGCTGGAAGGAGACGCAGCT 873

Qy       988 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 1047
             |||||||||||||||| || || ||| || |||| |||| |||||| || | || |||
Db       874 TTGAACAAATTCTTCCGGGAGATATACCAGAATGCGGATGAGGATATGAGGCGAGCGATG 933

Qy      1048 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1107
             |  || ||||| |||||||| ||||| || || || ||||| || ||| |||| |||||
Db       934 AGCAAATCATTTGTGGAATCGAATGGGACAGTGCTGTCAACAAACTGGCAAGAGGTTGGG 993

Qy      1108 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1167
               | ||||   | ||  | |  || || ||||| |||||||||||||| |||||   || |
Db       994 ACTAAGACAATCGAGAGTACTCCTCCGGATGGCATGGAGCTCAAGAAATGGGAGATCTGA 1053
```

Search completed: April 23, 2005, 08:51:26
Job time : 6464.31 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 22, 2005, 23:55:26 ; Search time 628.264 Seconds
(without alignments)
10232.696 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcggatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	733.6	67.6	1104	12	ADQ37150	Adq37150 Cell prol
	2	733.6	67.6	1104	12	ADQ15734	Adq15734 Rice stre
	3	490	45.1	571	10	ADF51379	Adf51379 Maize Rar
	4	419.2	38.6	1366	3	AAC39274	Aac39274 Arabidops
	5	418.8	38.6	1324	3	AAC40498	Aac40498 Arabidops
	6	417.6	38.5	1077	6	ADG87792	Adg87792 A. thalia
	7	417.6	38.5	1361	3	AAC51074	Aac51074 Arabidops
	8	393.2	36.2	1117	3	AAC42413	Aac42413 Arabidops
	9	286	26.3	287	6	ABL71197	Abl71197 Corn tass
c	10	231.4	21.3	581	9	ACL25190	Acl25190 DNA clone
	11	221.8	20.4	585	13	ACN47711	Acn47711 Cotton pr
	12	217.8	20.1	534	13	ACN59636	Acn59636 Cotton gy
	13	216.4	19.9	272	6	ABL75330	Abl75330 Corn tass
c	14	205.6	18.9	541	9	ACL25189	Acl25189 DNA clone
c	15	194.8	17.9	571	9	ACL25186	Acl25186 DNA clone
	16	194.4	17.9	528	13	ACN59570	Acn59570 Cotton gy
c	17	183.6	16.9	573	13	ACN47765	Acn47765 Cotton pr
c	18	180.8	16.6	500	9	ACL25185	Acl25185 DNA clone
c	19	175.4	16.2	523	13	ACN61180	Acn61180 Cotton gy
	20	164.4	15.1	272	10	ABX86113	Abx86113 Corn ear-
	21	156.4	14.4	1025	3	AAZ52451	Aaz52451 HTRM clon
	22	156.4	14.4	1185	6	ABL42163	Abl42163 Nucleotid
	23	156.4	14.4	1264	6	ABQ55038	Abq55038 Human ova
	24	156.4	14.4	1366	8	ABZ24700	Abz24700 Human cel
	25	156.4	14.4	1535	8	ACD13363	Acd13363 Human DNA
	26	156.4	14.4	1555	10	ADD29650	Add29650 Human tum
	27	156.4	14.4	1776	5	ABV28161	Abv28161 Human pro
	28	156.4	14.4	1776	5	ABV22337	Abv22337 Human pro
	29	156.4	14.4	2053	2	AAX27345	Aax27345 Human sec
	30	156.4	14.4	2053	9	ADA07224	Ada07224 Human cDN
	31	156.4	14.4	2053	12	ADN40960	Adn40960 Novel hum
	32	148.6	13.7	291	6	ABL73761	Abl73761 Corn tass
c	33	141.4	13.0	409	8	ABX62506	Abx62506 Arabidops
	34	133.6	12.3	615	13	ADQ55550	Adq55550 Novel can
	35	129.2	11.9	481	10	ADE82158	Ade82158 Arabidops
	36	125	11.5	388	10	ADF51380	Adf51380 Maize Rar
c	37	111.4	10.3	464	10	ADB50274	Adb50274 Primary r
c	38	102.4	9.4	340	9	ACL25187	Acl25187 DNA clone
c	39	102.4	9.4	412	9	ACL25184	Acl25184 DNA clone
	40	93.2	8.6	672	12	ADJ40154	Adj40154 Plant cDN
	41	90.8	8.4	1242	6	ABZ31779	Abz31779 Candida a
	42	78.8	7.3	431	6	ABL62696	Abl62696 Colon ade
	43	78	7.2	1303	12	ADJ43595	Adj43595 Plant cDN
	44	75.8	7.0	518	6	ABQ55648	Abq55648 Human ova
	45	75.8	7.0	869	4	AAK63551	Aak63551 Human imm

ALIGNMENTS

RESULT 1
ADQ37150
ID ADQ37150 standard; DNA; 1104 BP.
XX
AC ADQ37150;
XX
DT 07-OCT-2004 (first entry)
XX
DE Cell proliferation-related nucleic acid sequence #155.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response; ds.
XX
OS Oryza sativa.
XX
PN WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
PS Claim 57; SEQ ID NO 309; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body

CC of the specification as EPO data.
XX
SQ Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

Query Match 67.6%; Score 733.6; DB 12; Length 1104;
Best Local Similarity 80.9%; Pred. No. 6e-180;
Matches 888; Conservative 0; Mismatches 189; Indels 21; Gaps 2;

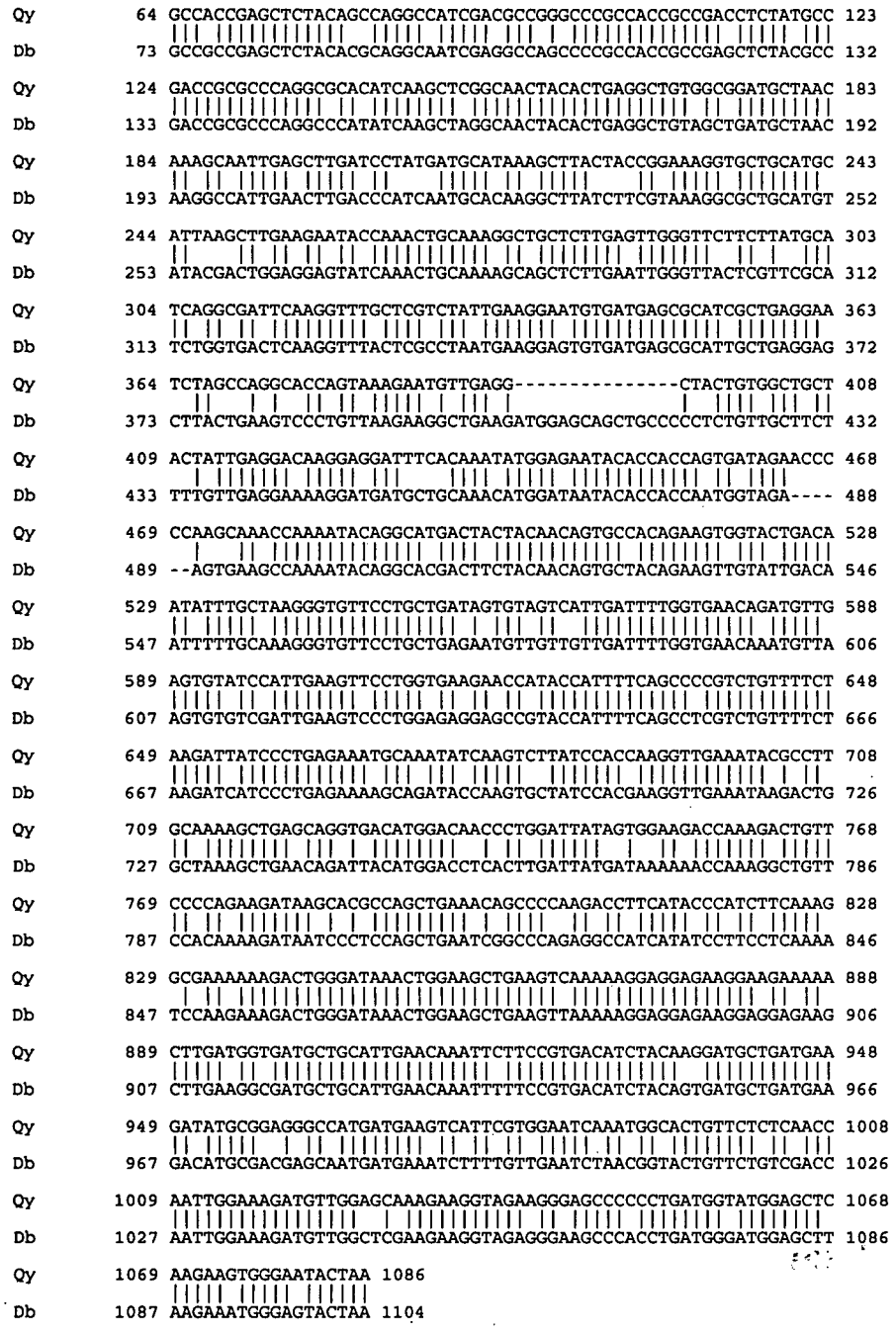
```
Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         13 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGACGACTTCGAGCTC 72

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||||  ||||| ||||| ||| | ||||||||||||||| |||||| |||
Db         73 GCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCCGAGCTCTACGCC 132

Qy        124 GACCGCGCCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              |||||||||||||| || |||||||| |||||||||||||||||||| || |||||||||
Db        133 GACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTAGCTGATGCTAAC 192

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              || || ||||| ||||| ||     ||||| || |||||    || ||||| ||||||||
Db        193 AAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAAGGCGCTGCATGT 252

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              ||    || || || ||||||||||| || |||||||| |||||||  || |  |||
Db        253 ATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTGGGTTACTCGTTCGCA 312

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              || || || ||||||||| |||| ||| ||||||| |||||||||||||| ||||||||
Db        313 TCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGCATTGCTGAGGAG 372

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGG----------------CTACTGTGGCTGCT 408
              ||   | |  || || ||||| | ||| |                |  |||| ||| ||
Db        373 CTTACTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCCCCCTCTGTTGCTTCT 432

Qy        409 ACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCC 468
              | ||||||| ||||| |||     |||| ||||| ||||||||||| || ||||
Db        433 TTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCAATGGTAGA---- 488

Qy        469 CCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACA 528
              |   || |||||||||||||| |||| |||||||||||| |||||||| ||| |||||
Db        489 --AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAAGTTGTATTGACA 546

Qy        529 ATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTG 588
              || ||||| |||||||||||||||||| | ||| ||  ||||||||||||||| |||||
Db        547 ATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGTGAACAAATGTTA 606

Qy        589 AGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCT 648
              ||||| || |||||||| ||||| || || || |||||||||||||| ||||||||||||
Db        607 AGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCTCGTCTGTTTTCT 666

Qy        649 AAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTT 708
              ||||| |||||||||||| ||| ||| |||||  ||||||| |||||||||||| | ||
Db        667 AAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTTGAAATAAGACTG 726

Qy        709 GCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTT 768
              || |||||||| ||| | |||||||  | || ||||||  |  || |||||||  |||||
Db        727 GCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAACCAAAGGCTGTT 786

Qy        769 CCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAG 828
              || || ||||||| | | ||||||||| | ||||  || || ||||| || || |||||
Db        787 CCACAAAAGATAATCCCTCCAGCTGAATCGGCCCAGAGGCCATCATATCCTTCCTCAAAA 846

Qy        829 GCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAA 888
              | || ||||||||||||||||||||||||||||| |||||||||||||||| || ||
Db        847 TCCAAGAAAGACTGGGATAAACTGGAAGCTGAAGTTAAAAAGGAGGAGAAGGAGGAGAAG 906

Qy        889 CTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAA 948
              ||||| || |||||||||||||||||||| |||||||||||||||  ||||||||||||
Db        907 CTTGAAGGCGATGCTGCATTGAACAAATTTTTCCGTGACATCTACAGTGATGCTGATGAA 966

Qy        949 GATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACC 1008
```

```
              || |||||  |.|| |||||||| || || || ||||| || || |||||||| || |||
Db        967 GACATGCGACGAGCAATGATGAAATCTTTTGTTGAATCTAACGGTACTGTTCTGTCGACC 1026

Qy       1009 AATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTC 1068
              |||||||||||||||||  | |||||||||| || ||||| |||||||| ||||||||
Db       1027 AATTGGAAAGATGTTGGCTCGAAGAAGGTAGAGGGAAGCCCACCTGATGGGATGGAGCTT 1086

Qy       1069 AAGAAGTGGGAATACTAA 1086
              ||||| ||||| ||||||
Db       1087 AAGAAATGGGAGTACTAA 1104
```

RESULT 2
ADQ15734
ID ADQ15734 standard; DNA; 1104 BP.
XX
AC ADQ15734;
XX
DT 07-OCT-2004 (first entry)
XX
DE Rice stress-related protein coding sequence #72.
XX
KW rice; stress-related protein; plant maturation; plant development;
KW plant proliferation; plant senescence; plant disease-resistance;
KW plant stress response; transgenic plant; pest tolerance;
KW herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
KW improved nutritional value; increased yield; increased proliferation;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2004061080-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041098.
XX
PR 26-DEC-2002; 2002US-0436564P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534374/51.
DR P-PSDB; ADQ15735.
XX
PT New isolated nucleic acids and proteins, useful for producing transgenic
PT plants having improved properties, e.g. tolerance to pests, herbicides,
PT or biotic or abiotic stresses, improved nutritional value, or increased
PT yield or proliferation.
XX
PS Claim 45; SEQ ID NO 143; 551pp; English.
XX
CC The invention comprises the amino acid and coding sequences of rice
CC stress-related proteins. The DNA and protein sequences of the invention
CC are useful for regulating and controlling plant maturation and
CC development, including proliferation, senescence, disease-resistance, or
CC stress response. They are also useful for producing transgenic plants
CC having improved properties, e.g. tolerance to pests, herbicides, or
CC biotic or abiotic stresses, improved nutritional value, increased yield
CC or proliferation, or improved structure causing less loss from lodging or
CC shattering. The present DNA sequence encodes a rice stress-related
CC protein of the invention.
XX
SQ Sequence 1104 BP; 323 A; 244 C; 284 G; 253 T; 0 U; 0 Other;

Query Match 67.6%; Score 733.6; DB 12; Length 1104;
Best Local Similarity 80.9%; Pred. No. 6e-180;
Matches 888; Conservative 0; Mismatches 189; Indels 21; Gaps 2;

```
Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         13 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGACGACTTCGAGCTC 72
```



RESULT 3
ADF51379
ID ADF51379 standard; cDNA; 571 BP.
XX
AC ADF51379;
XX
DT 12-FEB-2004 (first entry)
XX
DE Maize Rar interactior 2 (RarInt2) DNA seq id 2.
XX
KW plant protectant; gene therapy; Rar1; disease resistance; plant;
KW immunogen; maize; Rar interactor 2; RarInt2; gene; ss.
XX
OS Zea mays.
XX
PN US2003167504-A1.
XX
PD 04-SEP-2003.
XX
PF 12-FEB-2002; 2002US-00074473.
XX
PR 12-FEB-2001; 2001US-0268157P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Crane EH;
XX
DR WPI; 2003-898130/82.
XX
PT New maize Rar1-interactor nucleic acid useful for enhancing disease
PT resistance in plants, as a probe or amplification primer, in recombinant
PT expression of polypeptides, or as an immunogen in the preparation or
PT screening of antibodies.
XX
PS Claim 1; SEQ ID NO 2; 30pp; English.
XX
CC The invention describes an isolated nucleic acid capable of interacting
CC with Rar1. The nucleic acid comprises a member selected from: (a) a
CC polynucleotide comprising a sequence of 388 (S1), 571 (S2) or 691 (S3) bp
CC fully defined in the specification; (b) a polynucleotide amplified from a
CC Zea mays nucleic acid library using primers that selectively hybridise,
CC under high stringency conditions, to loci within the polynucleotide in
CC (a); (c) a polynucleotide encoding a maize Rar1-interactor protein; (d) a
CC polynucleotide having at least 80 or 90% sequence identity to (a); (e) a
CC polynucleotide which hybridises under high stringency conditions to (a);
CC and (f) a polynucleotide complementary to any of the above
CC polynucleotides. The nucleic acid molecule and methods are useful in
CC enhancing disease resistance in plants. The nucleic acid molecule may
CC also be used as a probe or an amplification primer, in recombinant
CC expression of polypeptides, or as an immunogen in the preparation and/or
CC screening of antibodies. This sequence represents a maize Rar interactor
CC 2 (RarInt2) polynucleotide.
XX
SQ Sequence 571 BP; 151 A; 154 C; 161 G; 105 T; 0 U; 0 Other;

Query Match 45.1%; Score 490; DB 10; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.1e-117;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

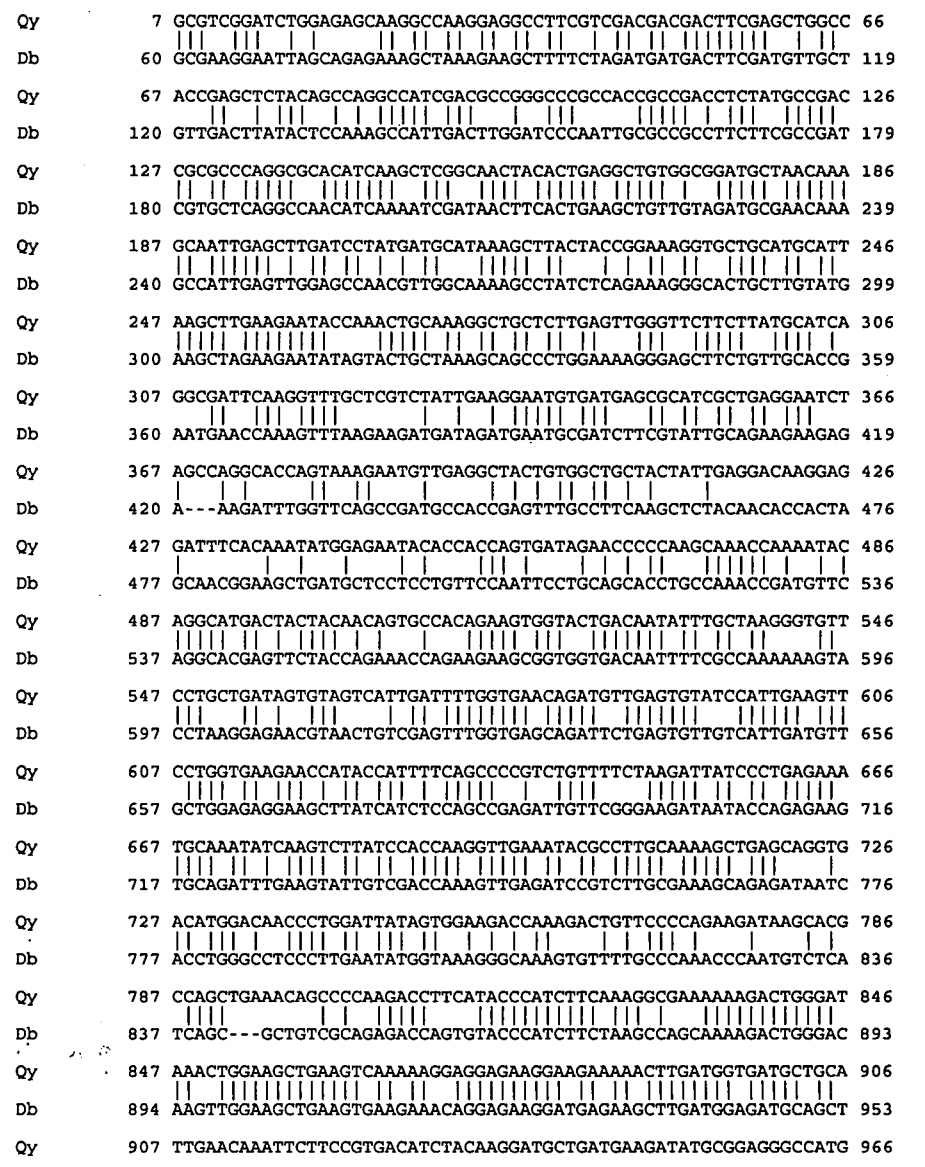
```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 321

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 381

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 441

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 501

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 561

Qy        481 AAATACAGGC 490
                ||||||||||
Db        562 AAATACAGGC 571
```

RESULT 4
AAC39274
ID AAC39274 standard; DNA; 1366 BP.
XX
AC AAC39274;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 24023.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.6%; Score 419.2; DB 3; Length 1366;
Best Local Similarity 63.1%; Pred. No. 2.9e-98;
Matches 681; Conservative 0; Mismatches 393; Indels 6; Gaps 2;



```
                 |||||||||| ||| | ||||| |||  |  ||| |||||||| ||| || |||| |||
Db        954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
              |  || ||||| |  || || ||||| || || || || || || |||||||| |||||
Db       1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
                | ||||| || ||  | |  || || ||||| ||||||||||||||||||||| || | |
Db       1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133


RESULT 5
AAC40498
ID   AAC40498 standard; DNA; 1324 BP.
XX
AC   AAC40498;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 28522.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
```

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.6%; Score 418.8; DB 3; Length 1324;
Best Local Similarity 63.0%; Pred. No. 3.6e-98;
Matches 680; Conservative 1; Mismatches 393; Indels 6; Gaps 2;

```
Qy        7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
            |||  |||  | |      || || || || || :|  | || || ||||||||  | ||
Db       60 GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTYTTCTAGATGATGACTTCGATGTTGCT 119

Qy       67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              ||  | |||  | | ||||| |||   |  |||     ||||| | |||  |||||
Db      120 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 179

Qy      127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
            || || |||||  |||||||  |||  |||| |||||| ||||| |  |||||  ||||||
Db      180 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 239

Qy      187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
            || |||||| | || || | | ||    ||||| ||    | | || ||  |||| || ||
Db      240 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 299

Qy      247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
            ||||| ||||||||     ||||| || || || || ||   |||  |||||  |||| |
Db      300 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 359

Qy      307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
              ||  ||| ||||        |  |  | ||||| |||   || || || || |||
Db      360 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 419

Qy      367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
            |   | |      ||  ||     |      | | | || || | |    |
Db      420 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476

Qy      427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
            |      |  |  |    | |    ||| |    | | | ||   ||||||  |  | |
Db      477 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 536

Qy      487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
            |||||  ||  | |||| | |    |  |||||  |||  |||||||  ||  ||  ||
Db      537 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 596

Qy      547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
            |||   ||  |  |||     | ||  |||||||| |||||  |||||||  ||||||  |||
Db      597 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 656

Qy      607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
            ||||  ||  |||  | ||  |||  | |||||   |  ||||     |||||  ||  ||  |||||
Db      657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 716

Qy      667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
            ||||  ||  |  ||||  ||  ||  |||||  |||||  ||  ||  |||||  |||||  |||     |
Db      717 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 776

Qy      727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
            ||  |||  |  ||||  ||  |||  ||  |  |  |  ||     |  |  |||  |     |  |
Db      777 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 836

Qy      787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
            ||||        |  |  |||||    ||||||||||| |||  |   ||||||||||||
Db      837 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893

Qy      847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
            ||  ||||||||||||||  ||  ||||||||||  ||  ||  ||||||||  |||||  ||
Db      894 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 953

Qy      907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
            ||||||||||  |||  |  |||||  |||   |  |||  ||||||||  |||  ||  ||||  |||
Db      954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy      967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
            |  ||  |||||  |   ||  ||  |||||  ||  ||  ||  ||  ||  ||  ||||||||  |||||
Db     1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy     1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
            |  |||||  ||  ||  | |  ||  ||  |||||  ||||||||||||||||||||  ||  | |
Db     1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133
```

```
RESULT 6
ADG87792
ID   ADG87792 standard; cDNA; 1077 BP.
XX
AC   ADG87792;
XX
DT   22-APR-2004  (first entry)
XX
DE   A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #234.
XX
KW   Pathogen infection-related gene; plant; Peronospora parasitica;
KW   defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW   oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200222675-A2.
XX
PD   21-MAR-2002.
XX
PF   14-SEP-2001; 2001WO-US028506.
XX
PR   15-SEP-2000; 2000US-0232778P.
PR   22-JUN-2001; 2001US-0300183P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (UYNC-) UNIV NORTH CAROLINA.
PA   (GLAZ/) GLAZEBROOK J.
PA   (WANG/) WANG X.
PA   (DANG/) DANGL J L.
PA   (EULG/) EULGEM T.
PA   (ZHUT/) ZHU T.
XX
PI   Glazebrook J,  Wang X,  Dangl JL,  Eulgem T,  Zhu T;
XX
DR   WPI; 2002-292409/33.
XX
PT   Novel isolated polynucleotide, useful for conveying pathogen resistance
PT   to plants, and for identifying plants infected with a pathogen.
XX
PS   Claim 3; SEQ ID NO 234; 605pp; English.
XX
CC   The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC   ADG87557}) whose expression is altered in response to pathogen infection,
CC   and to homologues of these genes from other plants or fungi, especially
CC   from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC   cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC   expression of genes of the invention was upregulated or downregulated in
CC   Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC   indicating that they play a role in defence mechanisms. The genes of the
CC   invention are regulated by RPP7 or RRP8 which act via unconventional
CC   signalling cascades, or by the RPP4-dependent pathway. The invention also
CC   relates to polypeptides encoded by the pathogen infection-related genes;
CC   promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC   ; expression cassettes, host cells and pathogen-resistant transgenic
CC   plants and their progeny comprising a polynucleotide of the invention;
CC   and a method of identifying a plant cell infected with a pathogen. The
CC   polynucleotide sequences and methods of the invention are useful for
CC   identifying plants infected with a pathogen, and for conferring
CC   resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC   neamotodes and insects (e.g., aphids). The present sequence represents an
CC   Arabidopsis thaliana gene whose expression is altered in response to
CC   Peronospora parasitica infection. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1077 BP; 337 A; 219 C; 277 G; 244 T; 0 U; 0 Other;

  Query Match           38.5%;  Score 417.6;  DB 6;  Length 1077;
  Best Local Similarity  63.0%;  Pred. No. 6.8e-98;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy        7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db        4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy       67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db       64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy      127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db      124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy      187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db      184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy      247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db      244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy      307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db      304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy      367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db      364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy      427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db      421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy      487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db      481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy      547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db      541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy      607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db      601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy      667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db      661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy      727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db      721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy      787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db      781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy      847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db      838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy      907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db      898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy      967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db      958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy     1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db     1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077
```

RESULT 7
AAC51074
ID AAC51074 standard; DNA; 1361 BP.
XX
AC AAC51074;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67179.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.5%; Score 417.6; DB 3; Length 1361;
Best Local Similarity 63.0%; Pred. No. 7.5e-98;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

RESULT 8
AAC42413
ID AAC42413 standard; DNA; 1117 BP.
XX
AC AAC42413;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35460.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

```
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match           36.2%;  Score 393.2;  DB 3;  Length 1117;
  Best Local Similarity   62.7%;  Pred. No. 1.5e-91;
  Matches  677;  Conservative    0;  Mismatches  373;  Indels   30;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              |||  ||| || |     ||||| || ||||| ||||| ||||| ||||||||  | ||
Db         68 GCGAAGGAGCTTGCTGATAAGGCTAAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT 127

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
                ||  | |||  | | ||||| |||   |  ||       ||| ||  |||  || ||
Db        128 GTTGACTTGTACTCCAAAGCCATTGACTTGGATCCTAATTGCGCTGAGTTCTTCGCTGAT 187

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              || || |||||  | ||||| ||||  | || |||||||| ||||| ||||| ||||||
Db        188 CGTGCTCAGGCCTATATCAAACTCGAAAGCTTCACTGAGGCCGTGGCAGATGCGAACAAA 247

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              |||||||||  | ||||||    ||   |||||||||     | |||||  |||| || ||
Db        248 GCAATTGAGTTGGATCCTTCATTGACCAAAGCTTACCTAAGAAAAGGAACTGCCTGTATG 307

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
              ||||||||||  || |  ||||| ||  | |||||||    |||| | |||     | |
Db        308 AAGCTTGAAGAGTATCGGACTGCTAAAACAGCTCTTGAAAAGGGTGCCTCTATCACGCCG 367

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
              | || || |  |||       |  |  ||||  ||   |  ||| | || |||
Db        368 AGTGAATCCAAATTTAAGAAGTTGATAGATGAATGCAATTTTCTAATCACAGAAGAA--- 424

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
                          | | |   |||  | || |  ||       |||   ||||
Db        425 ------------GAGAAAGATTTGGTTCAACCGGTGCCTTCGACTTTGCCTTCAAGTGT 471

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
              ||   ||| |              ||  | | |   | ||   | |    | | | || |||
Db        472 GACAGCACCA-----------CCAGTATCTGAACTTGATGTCACCCCTACAGCCAAGTAC 520

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
              ||||| || |||||| | |   |     ||||| ||  | ||  | ||||| || ||  |
Db        521 AGGCACGAGTACTACCAAAAGCCAGAGGAAGTTGTGGTAACTGTTTTTGCAAAAGGAATA 580

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
              ||     | | |||   ||| || |||||||||| ||  |||||||    ||||| |||
Db        581 CCCAAGCAGAATGTTAACATCGACTTTGGTGAACAAATTCTGAGTGTTGTGATTGAGGTT 640

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
              ||||| || ||  | ||  || | || ||  |  |||||   ||||| || || || ||
Db        641 CCTGGAGAGGATGCGTATTATCTCCAACCGAGATTGTTTGGAAAGATAATACCAGATAAG 700

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
              |||||||||  |||| || || |||||  |||| ||  | ||||||||||  ||     |
Db        701 TGCAAATATGAAGTATTGTCGACCAAAATTGAGATCTGTCTTGCAAAAGCCGATATAATC 760

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
              ||||||  |  |||| ||  |  |  | | |||  |  | | ||  ||   |   | ||
Db        761 ACATGGGCCTCCCTCGAACACGGCAAAGGGCCAGCGGTTTTGCCAAAGCCTAATGTCTCA 820

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
                |||  |   ||   |  || ||  | ||.||||||| |||      || ||||||||
Db        821 TCAGAGGTTTCA---CAGAGGCCAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGGAC 877

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
              ||  |||||||||||||| || ||| |||| || || || || || || || || || ||
Db        878 AAGCTGGAAGCTGAAGTGAAGAAGCAGGAAAAAGATGAGAAGCTGGAAGGAGACGCAGCT 937

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
              ||||||||||||||||| || || ||| || |||| ||||| |||||| || | || |||
Db        938 TTGAACAAATTCTTCCGGGAGATATACCAGAATGCGGATGAGGATATGAGGCGAGCGATG 997

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
              |  || ||||| |||||||| ||||| || || || ||||| || ||| |||| |||||
Db        998 AGCAAATCATTTGTGGAATCGAATGGGACAGTGCTGTCAACAAACTGGCAAGAGGTTGGG 1057

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              | ||||   | ||  | |  || || ||||| ||||||||||||||| |||||   || |
Db       1058 ACTAAGACAATCGAGAGTACTCCTCCGGATGGCATGGAGCTCAAGAAATGGGAGATCTGA 1117
```

```
RESULT 9
ABL71197
ID   ABL71197 standard; cDNA; 287 BP.
XX
AC   ABL71197;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:571.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
```

```
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 571; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC   hybridisation or amplification technologies, as query sequences to
CC   determine homology of known sequences, as probe for use in Southern or
CC   Northern hybridisation, and to identify the presence of and/or to
CC   determine the degree of similarity between two (or more) nucleic acid
CC   sequences
XX
SQ   Sequence 287 BP; 89 A; 60 C; 70 G; 68 T; 0 U; 0 Other;

  Query Match             26.3%;  Score 286;  DB 6;  Length 287;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-64;
  Matches  286;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 AGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGC 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGC 60

Qy        276 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAA 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAA 120

Qy        336 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGC 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGC 180

Qy        396 TACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACC 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACC 240

Qy        456 AGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 501
              |||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 286


RESULT 10
ACL25190/c
ID   ACL25190 standard; DNA; 581 BP.
XX
AC   ACL25190;
XX
DT   27-OCT-2003  (revised)
DT   17-OCT-2003  (first entry)
XX
DE   DNA clone originating in barley containing SNP encoding sequence #15181.
XX
KW   Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW   gene; ss.
XX
OS   Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN   WO2003057877-A1.
XX
PD   17-JUL-2003.
XX
PF   16-DEC-2002; 2002WO-IB005403.
XX
PR   20-DEC-2001; 2001JP-00387059.
PR   20-DEC-2001; 2001JP-00387131.
PR   20-DEC-2001; 2001JP-00403299.
PR   20-DEC-2001; 2001JP-00403300.
PR   27-SEP-2002; 2002JP-00327515.
XX
PA   (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI   Sato K,  Takeda K,  Kohara Y;
XX
DR   WPI; 2003-587127/55.
XX
PT   Single nucleotide polymorphism sites in barley varieties and DNA
PT   sequences containing them for analysis and identification of barley
PT   varieties and production of barley transformants with desired
PT   characteristics.
XX
PS   Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC   The present invention relates to oligonucleotide clones originating in
CC   barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC   (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC   varieties, identification of particular varieties and genotype-phenotype
CC   analysis, isolation of specific genes and creation of new varieties by
CC   transformation of barley varieties with them and production of new barley
CC   varieties with desired properties. The present sequence represents an
CC   oligonucleotide clone DNA sequence featured in the specification. The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 581 BP; 157 A; 151 C; 101 G; 172 T; 0 U; 0 Other;

  Query Match             21.3%;  Score 231.4;  DB 9;  Length 581;
  Best Local Similarity   82.6%;  Pred. No. 1.2e-49;
  Matches  265;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy        766 GTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCA 825
              | ||| ||||||||||      ||||||||| ||||||  || || || || || || ||| 
Db        569 GCTCCTCAGAAGATAAATGTACCAGCTGAATCAGCCCAGAGGCCATCTTATCCTTCATCA 510

Qy        826 AAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAA 885
              ||  | ||||| |||||||||||| || || |||||||| |||||  |||||||||| ||
Db        509 AAATCCAAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAACAGGAGAAGGATGAG 450

Qy        886 AAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGAT 945
              |||||||| |||||||||||||||||||| || |||||||| || ||||  |||||||||
Db        449 AAACTTGACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATTTACAGTGATGCTGAT 390

Qy        946 GAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCA 1005
              |||||||||| || || |||||||||| || ||||| || ||||| || |||||||||
Db        389 GAAGATATGCGTAGAGCAATGATGAAGTCCTTTGTGGAGTCTAATGGAACCGTTCTCTCA 330

Qy       1006 ACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCCTGATGGTATGGAG 1065
              ||||| |||||||||||| ||   ||||| ||| ||||| ||||| |||||||| ||||||
Db        329 ACCAACTGGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCTCCTGATGGAATGGAG 270

Qy       1066 CTCAAGAAGTGGGAATACTAA 1086
              |||||||||||||| || |||
```

```
Db        269 CTCAAGAAGTGGGAGTATTAA 249


RESULT 11
ACN47711
ID   ACN47711 standard; cDNA; 585 BP.
XX
AC   ACN47711;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Cotton primed seed EST Clone ID: LIB3825-014-Q1-K6-D9, SEQ:2492.
XX
KW   Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
KW   variety DP50B; library LIB3825; molecular tag; molecular marker;
KW   genetic mapping; molecular mapping; seed germination; plant growth;
KW   plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS   Gossypium hirsutum.
XX
PN   US2004123340-A1.
XX
PD   24-JUN-2004.
XX
PF   12-DEC-2001; 2001US-00021323.
XX
PR   14-DEC-2000; 2000US-0255619P.
XX
PA   (DEIK/) DEIKMAN J.
PA   (FENG/) FENG P C C.
PA   (FINC/) FINCHER K L.
PA   (ZIEG/) ZIEGLER T E.
XX
PI   Deikman J,  Feng PCC,  Fincher KL,  Ziegler TE;
XX
DR   WPI; 2004-479808/45.
XX
PT   New isolated nucleic acid molecule that encodes a plant protein or its
PT   fragment, useful for isolating a variety of agronomically significant
PT   genes associated with plant growth, quality or yield, and as molecular
PT   tags to map genes.
XX
PS   Claim 1; SEQ ID NO 2492; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 585 BP; 186 A; 117 C; 134 G; 148 T; 0 U; 0 Other;

  Query Match             20.4%;  Score 221.8;  DB 13;  Length 585;
  Best Local Similarity   62.7%;  Pred. No. 3.6e-47;
  Matches  362;  Conservative    0;  Mismatches  212;  Indels    3;  Gaps    1;

Qy        148 CTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATG 207
              |||  |||    | ||||  ||||||||  ||  ||||||||||||||||||  | |||||
Db          1 CTCAACAATCTCCCTGACGCTGTGGCAGAAGCTAACAAAGCAATTGAGTTGGATCCGTCC 60

Qy        208 ATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACT 267
              |||  ||||  |  ||||  ||  |||| |  ||||  ||  ||  ||||||||  ||  ||||||
Db         61 ATGTCTAAATCCTACTTGCGTAAAGCTACTGCGTGTATGAAGCTTGAGGAGTATCAAACT 120

Qy        268 GCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGT 327
              ||  ||||||||  |  |||   ||  ||  |||  |||  |||  ||  ||  ||  |||  |
Db        121 GCTAAGGCTGCGTTGGAGACTGGGGCTGCTTTGGCACCAGAAGACTCGAGATTTTCCAAG 180

Qy        328 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAT 387
              |  |  ||  ||||||||  ||||||||  ||  ||||||  |  |  ||  ||||    |||
Db        181 TTGATTAAAGAATGTGAAGAGCGCATTGCAGAGGAAAATGGTGAGTTACCAAAGGAGACG 240

Qy        388 GTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAAT 447
              |  ||  |           |   |  |  |   ||  ||  |   |   |    |||  |
Db        241 TTGGAAGAGGTGCCAACAAATGTTGTACCTGCGAAAGACATCTCTTCTGTCCAGGATATA 300

Qy        448 ACACCACCAGTGA---TAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAAC 504
              |    ||  |||   |  |  |  |||  ||||  |||  ||||||||||  |  ||  |
Db        301 CCTGATCCCATGACTGTTGCAGCACCACCCAAATCAACTTACAGGCATGAATTTTATCAG 360

Qy        505 AGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTC 564
              |   |    ||||||||  |  ||||||||||| |||||  |  ||    ||  ||||
Db        361 AAACCAGAGGAAGTGGTTGTCACAATATTTGCCAAGGGAATACCACGCGAGTGTGTTAAA 420

Qy        565 ATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATAC 624
              ||  |||  |||||||||  ||  |  |||||  |  |  |  |  ||  ||  ||||  | ||
Db        421 GTTAATTATGGTGAACAAATACTAAGTGTTGCTGTCAATGCACCCGGCAAAGATGCTTAT 480

Qy        625 CATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTA 684
              |||  |  ||  ||  ||  |  |||   |||||  ||  |||||  ||  ||||  ||||  |  ||  ||
Db        481 CATCTCCAACCTCGCTTATTTGGAAAGATAATACCTGACAAGTGCAGATATGATGTTTTG 540

Qy        685 TCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGC 721
              ||  |||||  |||||||  |  |  ||||||||||  |
Db        541 TCAACCAAAATTGAAATTAGGTTAGCAAAAGCTGAAC 577


RESULT 12
ACN59636
ID   ACN59636 standard; cDNA; 534 BP.
XX
AC   ACN59636;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Cotton gynoecium tissue EST Clone ID: LIB3829-033-Q1-K6-G10, SEQ:14417.
XX
KW   Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW   variety Nucotton33B; library LIB3829; molecular tag; molecular marker;
KW   genetic mapping; molecular mapping; seed germination; plant growth;
KW   plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS   Gossypium hirsutum.
XX
PN   US2004123340-A1.
XX
PD   24-JUN-2004.
XX
PF   12-DEC-2001; 2001US-00021323.
XX
PR   14-DEC-2000; 2000US-0255619P.
```

```
XX
PA   (DEIK/) DEIKMAN J.
PA   (FENG/) FENG P C C.
PA   (FINC/) FINCHER K L.
PA   (ZIEG/) ZIEGLER T E.
XX
PI   Deikman J,  Feng PCC,  Fincher KL,  Ziegler TE;
XX
DR   WPI; 2004-479808/45.
XX
PT   New isolated nucleic acid molecule that encodes a plant protein or its
PT   fragment, useful for isolating a variety of agronomically significant
PT   genes associated with plant growth, quality or yield, and as molecular
PT   tags to map genes.
XX
PS   Claim 1; SEQ ID NO 14417; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the US
CC   patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 534 BP; 181 A; 105 C; 115 G; 133 T; 0 U; 0 Other;

  Query Match           20.1%;  Score 217.8;  DB 13;  Length 534;
  Best Local Similarity   65.2%;  Pred. No. 3.8e-46;
  Matches  337;  Conservative    0;  Mismatches  177;  Indels    3;  Gaps    1;

Qy        461 TAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGG 520
              | | | | |||  |||| |||  |||||||||||| |  ||  | |    |    |||||||
Db         17 TTGCAGCACCACCCAAATCAACTTACAGGCATGAATTTTATCAGAAACCAGAGGAAGTGG 76

Qy        521 TACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAAC 580
              |  | ||||||||||| |||||  | ||    ||  ||||     || ||| ||||||||
Db         77 TTGTCACAATATTTGCCAAGGGAATACCACGCGAGTGTGTTAAAGTTAATTATGGTGAAC 136

Qy        581 AGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTC 640
              | ||  | |||||  |  |  | |  || ||  ||||  | || ||| | || || ||
Db        137 AAATACTAAGTGTTGCTGTCAATGCACCCGGCAAAGATGCTTATCATCTCCAACCTCGCT 196

Qy        641 TGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAA 700
              | |||   ||||| || ||||| || |||| |||| | || || || |||||  ||||||
Db        197 TATTTGGAAAGATAATACCTGACAAGTGCAGATATGATGTTTTGTCAACCAAAATTGAAA 256

Qy        701 TACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAA 760
              |  |  | ||||||||||| |   |   ||||||| |||| || | |||      |
Db        257 TTAGGTTAGCAAAAGCTGAACCAATTCAATGGACATCCCTCGAATTTAGCATGGAAGTTG 316

Qy        761 AGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCAT 820
              |     || |       |  | | ||   ||    | ||||| || |||||||
Db        317 CTGTGCCCCAAAGGATTAATGTATCTTCTGTTGCTGCAAATCAAAGACCATCGTACCCAT 376

Qy        821 CTTCAAAGGCGAAAA---AAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGA 877
              | || ||  | ||||     || |||||||| | ||||| |||| ||||||||||||||
Db        377 CCTCCAAACCAAAAAGGGTTGATTGGGATAAAATCGAAGCTCAAGTGAAAAAGGAGGAGA 436

Qy        878 AGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGG 937
              | || ||||| || |||||||||| || ||||| ||||| ||||| ||||| ||  |||
Db        437 AAGATGAAAAGCTAGATGGTGATGCAGCTTTGAATAAATTTTTCCGCGACATTTATCAGG 496

Qy        938 ATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTC 974
              ||||||| || |||    | |||||||||  ||| ||
Db        497 ATGCTGACGAGGATGCAAGAAGGGCCATGCAGAAATC 533


RESULT 13
ABL75330
ID   ABL75330 standard; cDNA; 272 BP.
XX
AC   ABL75330;
XX
DT   14-MAY-2002  (first entry)
XX
DE   Corn tassel-derived polynucleotide (cdps) SEQ ID NO:4704.
XX
KW   Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW   inheritance; characteristic; growth; development; disease resistance;
KW   environmental adaptability; quality; yield; molecular marker;
KW   multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS   Zea mays.
XX
PN   US2001051335-A1.
XX
PD   13-DEC-2001.
XX
PF   16-APR-1999;   99US-00294093.
XX
PR   21-APR-1998;   98US-0082567P.
XX
PA   (LALG/) LALGUDI R V.
PA   (ITOL/) ITO L Y.
PA   (SHER/) SHERMAN B K.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2002-163647/21.
XX
PT   Novel purified corn tassel-derived polynucleotide useful for determining
PT   altered gene expression, to recover regulatory elements and to follow
PT   inheritance of desirable characteristics through hybrid breeding
PT   programs.
XX
PS   Claim 1; SEQ ID NO 4704; 201pp; English.
XX
CC   The present sequence describes a purified corn tassel-derived
CC   polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC   selected from those given in ABL70627 to ABL76833. The cdps sequences
CC   encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC   can be used for determining altered gene expression, to recover
CC   regulatory elements and to follow inheritance of desirable
CC   characteristics through hybrid breeding programs. (I) are also useful in
CC   the evaluation, and alteration of desired characteristics associated with
CC   growth and development, disease resistance, environmental adaptability,
CC   quality and yield, and as molecular markers for studying inheritance of
CC   multigene traits in a plant breeding program. (I) can be used to produce
CC   a tassel-specific profile of gene transcription, a transcript image, to
CC   clone regulatory elements for use in transformation vectors, to express a
CC   polypeptide, to identify, isolate or extend identical or related corn
CC   tassel nucleic acid sequences from DNA libraries, in nucleic acid
```

CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences
XX
SQ Sequence 272 BP; 95 A; 54 C; 70 G; 49 T; 0 U; 4 Other;

Query Match 19.9%; Score 216.4; DB 6; Length 272;
Best Local Similarity 90.9%; Pred. No. 6.6e-46;
Matches 241; Conservative 0; Mismatches 19; Indels 5; Gaps 1;

```
Qy        703 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 60

Qy        763 ACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCT 822
              || |||||||||||||||||||||||| |||     | || ||||||||||||||||||
Db         61 GCTATTCCCCAGAAGATAAGCACGCCAGCNGAA-----CACANGACCTTCATACCCATCT 115

Qy        823 TCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAA 882
              |||||| | ||||| ||||||||||||||||||||||||| ||||||||||||||||||
Db        116 TCAAAGTCAAAAAAGGACTGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAA 175

Qy        883 GAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCT 942
              |||||||||| |||||||||||||||||||| |||||||||  ||||||||||||||||
Db        176 GAAAAACTTGAAGGTGATGCTGCATTGAACAAGTTCTTCCGTGATCTCTACAAGGATGCT 235

Qy        943 GATGAAGATATGCGGAGGGCCATGA 967
              |||||||||||  |  ||||  ||
Db        236 GATGAAGATATGGCGGNGGCCTGGA 260
```

RESULT 14
ACL25189/c
ID ACL25189 standard; DNA; 541 BP.
XX
AC ACL25189;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #15180.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 541 BP; 151 A; 144 C; 89 G; 157 T; 0 U; 0 Other;

Query Match 18.9%; Score 205.6; DB 9; Length 541;
Best Local Similarity 84.1%; Pred. No. 5.7e-43;
Matches 232; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy        811 TCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAG 870
              || || || || |||||  | ||||| |||||||||| || || ||||||||| |||||
Db        539 TCTTATCCTTCATCAAAATCCAAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAA 480

Qy        871 GAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATC 930
              |||||||||| || |||||||| |||||||||||||||||||| || |||||||| ||
Db        479 CAGGAGAAGGATGAGAAACTTGACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATT 420

Qy        931 TACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAAT 990
              ||||  |||||||||||||||||||| || || |||||||||| || ||||| || |||
Db        419 TACAGTGATGCTGATGAAGATATGCGTAGAGCAATGATGAAGTCCTTTGTGGAGTCTAAT 360

Qy        991 GGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 1050
              || || |||||||||||||| |||||||||| ||   ||||| ||| ||||| |||||
Db        359 GGAACCGTTCTCTCAACCAACTGGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCT 300

Qy       1051 CCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              |||||||| |||||||||||||||||||| || |||
Db        299 CCTGATGGAATGGAGCTCAAGAAGTGGGAGTATTAA 264
```

RESULT 15
ACL25186/c
ID ACL25186 standard; DNA; 571 BP.
XX
AC ACL25186;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #15177.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX

DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 571 BP; 168 A; 148 C; 88 G; 166 T; 0 U; 1 Other;

Query Match 17.9%; Score 194.8; DB 9; Length 571;
Best Local Similarity 85.4%; Pred. No. 3.7e-40;
Matches 217; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy        833 AAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTG 892
              |||| |||||||||| || || |||||||| |||||  |||||||||| || |||||||
Db        563 AAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAACAGGAGAAGGATGAGAAACTTG 504

Qy        893 ATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATA 952
              | |||||||||||||||||| || |||||||| || ||||  ||||||||||||||||
Db        503 ACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATTTACAGTGATGCTGATGAAGATA 444

Qy        953 TGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATT 1012
              |||| || || |||||||||| || ||||| || ||||| || |||||||||||||| |
Db        443 TGCGTAGAGCAATGATGAAGTCCTTTGTGGAGTCTAATGGAACCGTTCTCTCAACCAACT 384

Qy       1013 GGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGA 1072
              |||||||||| ||   ||||| ||| ||||| ||||| |||||||| ||||||||||||
Db        383 GGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCTCCTGATGGAATGGAGCTCAAGA 324

Qy       1073 AGTGGGAATACTAA 1086
              ||||||| || |||
Db        323 AGTGGGAGTATTAA 310
```

Search completed: April 23, 2005, 03:11:00
Job time : 630.264 secs

This Page Blank (usptc)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:46:01 ; Search time 194.463 Seconds
(without alignments)
9137.976 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcggatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	164.4	15.1	272	4	US-09-313-294A-4573	Sequence 4573, Ap
2	156.4	14.4	2053	3	US-09-227-357-45	Sequence 45, Appl
3	76.2	7.0	198	4	US-09-248-796A-2330	Sequence 2330, Ap
4	74.2	6.8	1735	4	US-09-949-016-1499	Sequence 1499, Ap
5	74.2	6.8	1982	4	US-09-016-434-1067	Sequence 1067, Ap
6	74.2	6.8	1982	4	US-09-825-497A-40	Sequence 40, Appl
7	74.2	6.8	1987	4	US-09-825-497A-39	Sequence 39, Appl
8	74.2	6.8	1987	4	US-09-517-779-1	Sequence 1, Appli
9	73.2	6.7	425	4	US-09-513-999C-4027	Sequence 4027, Ap
10	63.4	5.8	47493	4	US-09-949-016-13241	Sequence 13241, A
11	62.2	5.7	989	4	US-09-270-767-12741	Sequence 12741, A
12	61.2	5.6	2735	4	US-09-551-974A-101	Sequence 101, App
13	61.2	5.6	2735	4	US-09-565-501A-101	Sequence 101, App
14	61.2	5.6	2735	4	US-09-639-206A-101	Sequence 101, App
15	61.2	5.6	2735	4	US-09-874-923-101	Sequence 101, App
16	61.2	5.6	3012	4	US-09-551-974A-94	Sequence 94, Appl
17	61.2	5.6	3012	4	US-09-565-501A-94	Sequence 94, Appl
18	61.2	5.6	3012	4	US-09-639-206A-94	Sequence 94, Appl
19	61.2	5.6	3012	4	US-09-874-923-94	Sequence 94, Appl
20	61.2	5.6	3134	2	US-08-533-669A-1	Sequence 1, Appli
21	61.2	5.6	3134	3	US-09-183-861-1	Sequence 1, Appli
22	61.2	5.6	3134	3	US-09-022-765-1	Sequence 1, Appli
23	61.2	5.6	3134	4	US-09-551-974A-1	Sequence 1, Appli
24	61.2	5.6	3134	4	US-09-565-501A-1	Sequence 1, Appli
25	61.2	5.6	3134	4	US-09-639-206A-1	Sequence 1, Appli
26	61.2	5.6	3134	4	US-09-874-923-1	Sequence 1, Appli
27	61.2	5.6	3134	4	US-08-798-841-1	Sequence 1, Appli
28	61.2	5.6	4233	4	US-09-551-974A-99	Sequence 99, Appl
29	61.2	5.6	4233	4	US-09-565-501A-99	Sequence 99, Appl
30	61.2	5.6	4233	4	US-09-639-206A-99	Sequence 99, Appl
31	61.2	5.6	4233	4	US-09-874-923-99	Sequence 99, Appl
32	61.2	5.6	4917	4	US-09-551-974A-100	Sequence 100, App
33	61.2	5.6	4917	4	US-09-565-501A-100	Sequence 100, App
34	61.2	5.6	4917	4	US-09-639-206A-100	Sequence 100, App
35	61.2	5.6	4917	4	US-09-874-923-100	Sequence 100, App
36	61.2	5.6	4929	4	US-09-551-974A-98	Sequence 98, Appl
37	61.2	5.6	4929	4	US-09-565-501A-98	Sequence 98, Appl
38	61.2	5.6	4929	4	US-09-639-206A-98	Sequence 98, Appl
39	61.2	5.6	4929	4	US-09-874-923-98	Sequence 98, Appl
c 40	58	5.3	7218	1	US-08-232-463-14	Sequence 14, Appl
41	56.4	5.2	2221	3	US-09-301-978C-1	Sequence 1, Appli
42	56.4	5.2	2229	4	US-09-949-016-1647	Sequence 1647, Ap
c 43	56.4	5.2	2274	4	US-09-220-132-188	Sequence 188, App
44	52.6	4.8	2111	4	US-09-949-016-1548	Sequence 1548, Ap
45	51.8	4.8	405	4	US-09-513-999C-11529	Sequence 11529, A

ALIGNMENTS

RESULT 1
US-09-313-294A-4573
; Sequence 4573, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4573
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348654H1
; NAME/KEY: unsure
; LOCATION: 92, 100, 103, 105, 112, 114, 140, 204, 210, 244-245, 250, 257, 261-262,
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4573

Query Match 15.1%; Score 164.4; DB 4; Length 272;
Best Local Similarity 77.3%; Pred. No. 8.7e-35;
Matches 214; Conservative 0; Mismatches 58; Indels 5; Gaps 2;

```
Qy      326 GTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGA 385
            ||||||||||||||||||||||||| |||| ||||||||||||||||||||||||| |||||
Db        1 GTCTATTGAAGGAATGTGATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGA 60

Qy      386 ATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGA 445
            |||| |||||| |||| ||||||||| |||| ||||||| || | || ||| | |||| |
Db       61 ATGTCGAGGCTCCTGTAGCTGCTACTGTTGANGACAAGGNGGNTNTCGCAANTNTGGATA 120

Qy      446 ATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACA 505
            ||||||| |||||| |||| |||||||||||||| ||||| ||||||||||||||||||
Db      121 ATACACCGCCAGTGGTAGANCCCCCAAGCAAACCTAAATATAGGCATGACTACTACAACA 180

Qy      506 GTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCA 565
            |||||||||||  |        ||     ||||  | ||    | | |||||||  ||||
Db      181 GTGCCACAGAATGG---TCCTCACATNTTGCCNAAGGGTGTTCCGCTGATAATGTA--TC 235

Qy      566 TTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGA 602
              |   ||   |||  ||||||  |||      ||  ||  |
Db      236 ATTGATTCNNTGATNAGATGTNGAGNNCNNCCCTTCA 272
```

```
RESULT 2
US-09-227-357-45
; Sequence 45, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P1
;  CURRENT APPLICATION NUMBER: US/09/227,357
;  CURRENT FILING DATE: 1999-01-08
;  EARLIER APPLICATION NUMBER: PCT/US98/13684
;  EARLIER FILING DATE: 1998-07-07
;  EARLIER APPLICATION NUMBER: 60/051,926
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,793
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,925
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,929
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,803
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,732
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,931
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,932
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,916
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,930
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,918
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,920
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,733
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,795
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,919
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,928
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/055,722
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,723
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,948
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,949
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,953
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,950
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,947
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,964
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/056,360
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,684
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,984
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,954
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/058,785
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,664
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,660
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,661
;  EARLIER FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 2053
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-227-357-45

  Query Match             14.4%;  Score 156.4;  DB 3;  Length 2053;
  Best Local Similarity   56.3%;  Pred. No. 3.7e-32;
  Matches  339;  Conservative    0;  Mismatches  251;  Indels   12;  Gaps    2;

Qy        479 CAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTA 538
              |||||  || | ||||||  ||  | |  |   | |||| ||  | ||| |  |    |
Db       1416 CAAAAATCAAGTATGACTGGTATCAAACAGAATCTCAAGTAGTCATTACACTTATGATCA 1475

Qy        539 AGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCA 598
              ||  |||||       ||  ||||    | || |||   ||| |   ||||  ||  |
Db       1476 AGAATGTTCAGAAGAATGATGTAAATGTGGAATTTTCAGAAAAAGAGTTGTCTGCTTTGG 1535

Qy        599 TTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCC 658
              || || |||||  || |||    ||| ||||  | |      ||  ||  |   || || |
Db       1536 TTAAACTTCCTTCTGGAGAGGATTACAATTTGAAACTGGAACTTCTTCATCCTATAATAC 1595

Qy        659 CTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTG 718
              | ||  |  |||  | | ||||  | || || ||| |||||||     ||   ||| | |
Db       1596 CAGAACAGAGCACGTTTAAAGTACTTTCAACAAAGATTGAAATTAAACTGAAAAAGCCAG 1655

Qy        719 AGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGA 778
              ||   |||| ||||  ||  || ||        ||   |  | |      | | |  ||
Db       1656 AGGCTGTGAGATGGGAAAAGCTAGA------GGGGCAAGGAGATGTGCCTACGCCAAAAC 1709

Qy        779 TAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAG 838
              |  |    ||| || ||    || |  || ||||   | |||        | || |
Db       1710 AATTCGTAGCAGATGTAAAGAACCTATATCCATCATCATCTCCTT------ATACAAGAA 1763

Qy        839 ACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTG 898
              | |||||||||| |||  | |||  ||||| | || || ||| | |||||  | || || |
Db       1764 ATTGGGATAAATTGGTTGGTGAGATCAAAGAAGAAGAAAAGAATGAAAAGTTGGAGGGAG 1823

Qy        899 ATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGA 958
              |||| || || |||| ||| || |   | |||||    |||| |  ||| ||  ||
Db       1824 ATGCAGCTTTAAACAGATTATTTCAGCAGATCTATTCAGATGGTTCTGATGAAGTGAAAC 1883

Qy        959 GGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAG 1018
              | |||||||  || || ||  |||| ||   ||| || ||| |    |||| |||   |
Db       1884 GTGCCATGAACAAATCCTTTATGGAGTCGGGTGGTACAGTTTTGAGTACCAACTGGTCTG 1943

Qy       1019 ATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGG 1078
              |||| ||   || ||| || |||   |  || ||||||| ||||||    || ||||
Db       1944 ATGTAGGTAAAAGGAAAGTTGAAATCAATCCTCCTGATGATATGGAATGGAAAAAGTACT 2003

Qy       1079 AA 1080
              ||
Db       2004 AA 2005


RESULT 3
US-09-248-796A-2330
; Sequence 2330, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
```

```
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2330
;    LENGTH: 198
;    TYPE: DNA
;    ORGANISM: Candida albicans
US-09-248-796A-2330

  Query Match             7.0%;  Score 76.2;  DB 4;  Length 198;
  Best Local Similarity   64.4%;  Pred. No. 6.9e-11;
  Matches  114;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy        910 AACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCCGGAGGGCCATGATG 969
              ||   | || |||||   | || | ||| |||| ||| ||||||     | || || ||||||
Db         22 AATGACTTTTTCCGGAAAATATTCAAAGATGTTGACGAAGATTCTAGAAGAGCGATGATG 81

Qy        970 AAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCA 1029
              |||    |  ||  |||| ||||| |||||| |   | || | |||| | || | |  ||
Db         82 AAGAGTTATGTTCAATCGAATGGAACTGTTTTGACTACAAGTTGGGATGAAGCTAAAGAT 141

Qy       1030 AAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              ||  |  | || |      || || ||||| |||||  ||||||| |||||    |||
Db        142 AAAGAATTTGAGGTTTTACCACCAGATGGGATGGAAGTCAAGAAATGGGATACGTAA 198


RESULT 4
US-09-949-016-1499
; Sequence 1499, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1499
;    LENGTH: 1735
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-1499

  Query Match             6.8%;  Score 74.2;  DB 4;  Length 1735;
  Best Local Similarity   55.1%;  Pred. No. 7.4e-10;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   |||   | | | |||| ||||   |||
Db         82 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 141

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||||     |||  ||  |||  |  ||||| | ||
Db        142 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 201

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |       |   |||     | ||  |||  || || |
Db        202 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 261

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || |||||||| ||     | |       || | ||| |||||    | |||       |||
Db        262 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 321

Qy        247 AAGCTTGAAGAATACCAAACTGC 269
```

```
              || |   | | ||   | ||
Db        322 GCACTGGGCAAGTTCCGGGCCGC 344


RESULT 5
US-09-016-434-1067
; Sequence 1067, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1067:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1982 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1122930
US-09-016-434-1067

  Query Match             6.8%;  Score 74.2;  DB 4;  Length 1982;
  Best Local Similarity   55.1%;  Pred. No. 7.9e-10;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   |||   | | | |||| ||||   |||
Db         60 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 119

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||     |||  ||  |||  |  ||||| | ||
Db        120 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 179

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |       |   |||     | ||  |||  || || |
Db        180 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 239

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || |||||||| ||     | |       || | ||| |||||    | |||       |||
Db        240 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 299

Qy        247 AAGCTTGAAGAATACCAAACTGC 269
```

```
                 || |   | | ||   | ||
Db        300 GCACTGGGCAAGTTCCGGGCCGC 322


RESULT 6
US-09-825-497A-40
; Sequence 40, Application US/09825497A
; Patent No. 6599742
; GENERAL INFORMATION:
;  APPLICANT: Honkanen, Richard  E.
;  APPLICANT:  Dean, Nicholas  M.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODUATION OF HUMAN SERINE/THREONINE PRC
;  TITLE OF INVENTION:  PHOSPHATASE GENE EXPRESSION
;  FILE REFERENCE: ISPH-0572
;  CURRENT APPLICATION NUMBER: US/09/825,497A
;  CURRENT FILING DATE:  2001-04-06
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
;   LENGTH: 1982
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-497A-40

  Query Match             6.8%;  Score 74.2;  DB 4;  Length 1982;
  Best Local Similarity   55.1%;  Pred. No. 7.9e-10;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         60 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 119

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        120 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 179

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |       |    |||   | ||  |||  || || |
Db        180 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 239

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || ||||||||| ||    | |      || | ||| |||||    | |||      |||
Db        240 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 299

Qy        247 AAGCTTGAAGAATACCAAACTGC 269
                || |   | | ||   | ||
Db        300 GCACTGGGCAAGTTCCGGGCCGC 322


RESULT 7
US-09-825-497A-39
; Sequence 39, Application US/09825497A
; Patent No. 6599742
; GENERAL INFORMATION:
;  APPLICANT: Honkanen, Richard  E.
;  APPLICANT:  Dean, Nicholas  M.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODUATION OF HUMAN SERINE/THREONINE PRC
;  TITLE OF INVENTION:  PHOSPHATASE GENE EXPRESSION
;  FILE REFERENCE: ISPH-0572
;  CURRENT APPLICATION NUMBER: US/09/825,497A
;  CURRENT FILING DATE:  2001-04-06
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-497A-39

  Query Match             6.8%;  Score 74.2;  DB 4;  Length 1987;
  Best Local Similarity   55.1%;  Pred. No. 7.9e-10;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         65 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 124

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        125 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 184

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |       |    |||   | ||  |||  || || |
Db        185 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 244

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || ||||||||| ||    | |      || | ||| |||||    | |||      |||
Db        245 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 304

Qy        247 AAGCTTGAAGAATACCAAACTGC 269
                || |   | | ||   | ||
Db        305 GCACTGGGCAAGTTCCGGGCCGC 327


RESULT 8
US-09-517-779-1
; Sequence 1, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
;  APPLICANT:  Luo, Ying
;  APPLICANT:  Xu, Xiang
;  TITLE OF INVENTION:  Cell Cycle Proteins Associated with Rad9, Compositions
;  TITLE OF INVENTION:  and Methods Of Use
;  FILE REFERENCE:  A68293/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/517,779
;  CURRENT FILING DATE:  2000-03-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-517-779-1

  Query Match             6.8%;  Score 74.2;  DB 4;  Length 1987;
  Best Local Similarity   55.1%;  Pred. No. 7.9e-10;
  Matches  145;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db         65 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 124

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        125 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 184

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |       |    |||   | ||  |||  || || |
Db        185 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 244

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || ||||||||| ||    | |      || | ||| |||||    | |||      |||
Db        245 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGCCAGCAACATG 304

Qy        247 AAGCTTGAAGAATACCAAACTGC 269
                || |   | | ||   | ||
Db        305 GCACTGGGCAAGTTCCGGGCCGC 327


RESULT 9
US-09-513-999C-4027
; Sequence 4027, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
```

```
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 4027
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 46..417
US-09-513-999C-4027

  Query Match             6.7%;  Score 73.2;  DB 4;  Length 425;
  Best Local Similarity   57.4%;  Pred. No. 6.7e-10;
  Matches  132;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   ||| ||  |||   ||||||| ||   ||||   | | | |||| ||||   |||
Db        127 GCAGAGGAGCTCAAGACTCAGGCCAATGACTACTTCAAAGCCAAGGACTACGAGAACGCC 186

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
              | | || ||||||||||||||||||||        |||  ||  |||  |  ||||| | ||
Db        187 ATCAAGTTCTACAGCCAGGCCATCGAGCTGAACCCCAGCAATGCCATCTACTATGGCAAC 246

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              |||  || |||  || |      |    |||     | ||  |||  || || |
Db        247 CGCAGCCTGGCCTACCTGCGCACTGAGTGCTATGGCTACGCGCTGGGAGACGCCACGCGG 306

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGC 236
              || |||||||| ||    | |      || | ||| |||||    | |||
Db        307 GCCATTGAGCTGGACAAGAAGTACATCAAGGGTTATTACCGCCGGGCTGC 356


RESULT 10
US-09-949-016-13241
; Sequence 13241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13241
;   LENGTH: 47493
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13241

  Query Match             5.8%;  Score 63.4;  DB 4;  Length 47493;
  Best Local Similarity   54.5%;  Pred. No. 3.5e-06;
  Matches  127;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy         37 GCCTTCGTCGACGACGACTTCGAGCTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCC 96
              | || ||  | | | |||| ||||   |||| | || ||||||||||||||||||||
Db       8641 GTCTCCGCAGCCAAGGACTACGAGAACGCCATCAAGTTCTACAGCCAGGCCATCGAGCTG 8700

Qy         97 GGGCCCGCCACCGCCGACCTCTATGCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAAC 156
              |||  ||  |||  |  ||||| | |||||  || |||  || |       |    |
Db       8701 AACCCCAGCAATGCCATCTACTATGGCAACCGCAGCCTGGCCTACCTGCGCACTGAGTGC 8760

Qy        157 TACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTATGATGCATAAA 216
              ||     | ||  |||  || || |     || |||||||| ||    | |       ||
Db       8761 TATGGCTACGCGCTGGGAGACGCCACGCGGGCCATTGAGCTGGACAAGAAGTACATCAAG 8820

Qy        217 GCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACTGC 269
              | ||| |||||    | |||      |||   || |   | | ||   | ||
Db       8821 GGTTATTACCGCCGGGCTGCCAGCAACATGGCACTGGGCAAGTTCCGGGCCGC 8873


RESULT 11
US-09-270-767-12741
; Sequence 12741, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12741
;   LENGTH: 989
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-12741

  Query Match             5.7%;  Score 62.2;  DB 4;  Length 989;
  Best Local Similarity   55.8%;  Pred. No. 1e-06;
  Matches  140;  Conservative    0;  Mismatches  108;  Indels    3;  Gaps    1;

Qy        835 AAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGAT 894
              ||  |||||||  | ||||      || |   |||||   ||  || | |   |      |
Db        446 AAGAACTGGGACCAGCTGGTCAGCGAGGAGGAAAAGATCGACGAGAAGG---AGGCCAAG 502

Qy        895 GGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATG 954
              || || || ||  | | |||  | |||    | |||||||       |       ||  |
Db        503 GGCGAGGCGGCCCTCACCAACCTGTTCAAGAAAATCTACAGCTCATCCTCGCCCGAGGTC 562

Qy        955 CGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGG 1014
              |  | ||||||||  ||||| |||  ||| ||    ||||| || |||   ||||| |||
Db        563 CAAAAGGCCATGAACAAGTCCTTCTCGGAGTCCGGAGGCACCGTGCTCAGCACCAACTGG 622

Qy       1015 AAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAG 1074
              || || || ||   | | |  ||     | ||  || || ||||  |  ||| ||    ||
Db        623 AATGAGGTGGGCAAAGAAAGAGTCACCGTGAAGCCTCCAAATGGAACCGAGTTCCGCGAG 682

Qy       1075 TGGGAATACTA 1085
              |||||  | ||
Db        683 TGGGAGAAGTA 693


RESULT 12
US-09-551-974A-101
; Sequence 101, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C5
;  CURRENT APPLICATION NUMBER: US/09/551,974A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-551-974A-101

  Query Match             5.6%;  Score 61.2;  DB 4;  Length 2735;
  Best Local Similarity   52.3%;  Pred. No. 3.2e-06;
  Matches  135;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          6 CGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGC 65
              |||| | |   | |   | | |  ||  ||  ||||   || ||  | |||     |||
Db       1698 CGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCAAGGAGGATAAGTTCCCCGAGGC 1757

Qy         66 CACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGA 125
              |   |    ||||   |||||||| |   |   || |||     |  |  |||   | |
Db       1758 CGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTGCCGAGCACACCTCCTACAGCAA 1817

Qy        126 CCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAA 185
              |||||    ||| ||||||||||| ||   || ||  || ||  |   ||| ||  | ||
Db       1818 TCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACGACGCCCTCAAGGACGCGGAGAA 1877

Qy        186 AGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCAT 245
              |||||||  | ||      |   ||| |  |||   || ||  |   ||| ||  | ||
Db       1878 GTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACGCGCGCAAGGGTCATGCTTACTT 1937

Qy        246 TAAGCTTGAAGAATACCA 263
              |  |   |  | ||| |
Db       1938 TTGGACCAAGCAGTACAA 1955


RESULT 13
US-09-565-501A-101
; Sequence 101, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Peter Probst
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C6
;  CURRENT APPLICATION NUMBER: US/09/565,501A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  112
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-565-501A-101

  Query Match             5.6%;  Score 61.2;  DB 4;  Length 2735;
  Best Local Similarity   52.3%;  Pred. No. 3.2e-06;
  Matches  135;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          6 CGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGC 65
              |||| | |   | |   | | |  ||  ||  ||||   || ||  | |||     |||
Db       1698 CGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCAAGGAGGATAAGTTCCCCGAGGC 1757

Qy         66 CACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGA 125
              |   |    ||||   |||||||| |   |   || |||     |  |  |||   | |
Db       1758 CGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTGCCGAGCACACCTCCTACAGCAA 1817

Qy        126 CCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAA 185
              |||||    ||| ||||||||||| ||   || ||  || ||  |   ||| ||  | ||
Db       1818 TCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACGACGCCCTCAAGGACGCGGAGAA 1877

Qy        186 AGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCAT 245
              |||||||  | ||      |   ||| |  |||   || ||  |   ||| ||  | ||
Db       1878 GTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACGCGCGCAAGGGTCATGCTTACTT 1937

Qy        246 TAAGCTTGAAGAATACCA 263
              |  |   |  | ||| |
Db       1938 TTGGACCAAGCAGTACAA 1955


RESULT 14
US-09-639-206A-101
; Sequence 101, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C7
;  CURRENT APPLICATION NUMBER: US/09/639,206A
;  CURRENT FILING DATE:  2000-08-14
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 101
;   LENGTH: 2735
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
;   OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-639-206A-101

  Query Match             5.6%;  Score 61.2;  DB 4;  Length 2735;
  Best Local Similarity   52.3%;  Pred. No. 3.2e-06;
  Matches  135;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy          6 CGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGC 65
              |||| | |   | |   | | |  ||  ||  ||||   || ||  | |||     |||
Db       1698 CGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCAAGGAGGATAAGTTCCCCGAGGC 1757

Qy         66 CACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGA 125
              |   |    ||||   |||||||| |   |   || |||     |  |  |||   | |
Db       1758 CGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTGCCGAGCACACCTCCTACAGCAA 1817

Qy        126 CCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAA 185
              |||||    ||| ||||||||||| ||   || ||  || ||  |   ||| ||  | ||
Db       1818 TCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACGACGCCCTCAAGGACGCGGAGAA 1877

Qy        186 AGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCAT 245
              |||||||  | ||      |   ||| |  |||   || ||  |   ||| ||  | ||
Db       1878 GTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACGCGCGCAAGGGTCATGCTTACTT 1937

Qy        246 TAAGCTTGAAGAATACCA 263
              |  |   |  | ||| |
Db       1938 TTGGACCAAGCAGTACAA 1955


RESULT 15
US-09-874-923-101
; Sequence 101, Application US/09874923
; Patent No. 6638517
```

; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence encoding fusion (poly-protein) constructs
; OTHER INFORMATION: comprising multiple Leishmania antigens
US-09-874-923-101

Query Match 5.6%; Score 61.2; DB 4; Length 2735;
Best Local Similarity 52.3%; Pred. No. 3.2e-06;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy          6 CGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGC 65
              ||||  | |   | |   | | | ||  ||  ||||   || ||  | |||     |||
Db       1698 CGCGAAGCAGAAGAAAGACGAAGGTAACCAGTACTTCAAGGAGGATAAGTTCCCCGAGGC 1757

Qy         66 CACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGA 125
              |   |    ||||   |||||||| |   |   || |||     |  |  |||   | |
Db       1758 CGTGGCAGCGTACACGGAGGCCATCAAGCGCAACCCTGCCGAGCACACCTCCTACAGCAA 1817

Qy        126 CCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAA 185
              |||||    ||| |||||||||| ||   || ||  || ||  |   ||| ||  | ||
Db       1818 TCGCGCGGCCGCGTACATCAAGCTTGGAGCCTTCAACGACGCCCTCAAGGACGCGGAGAA 1877

Qy        186 AGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCAT 245
              ||||||||  | ||     |   ||| |  |||   || || |||  ||| | | |
Db       1878 GTGCATTGAGCTGAAGCCCGACTTTGTTAAGGGCTACGCGCGCAAGGGTCATGCTTACTT 1937

Qy        246 TAAGCTTGAAGAATACCA 263
              |  |    |  | ||| |
Db       1938 TTGGACCAAGCAGTACAA 1955
```

Search completed: April 23, 2005, 08:59:04
Job time : 195.463 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 02:35:27 ; Search time 4819.69 Seconds
(without alignments)
10918.214 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcggatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

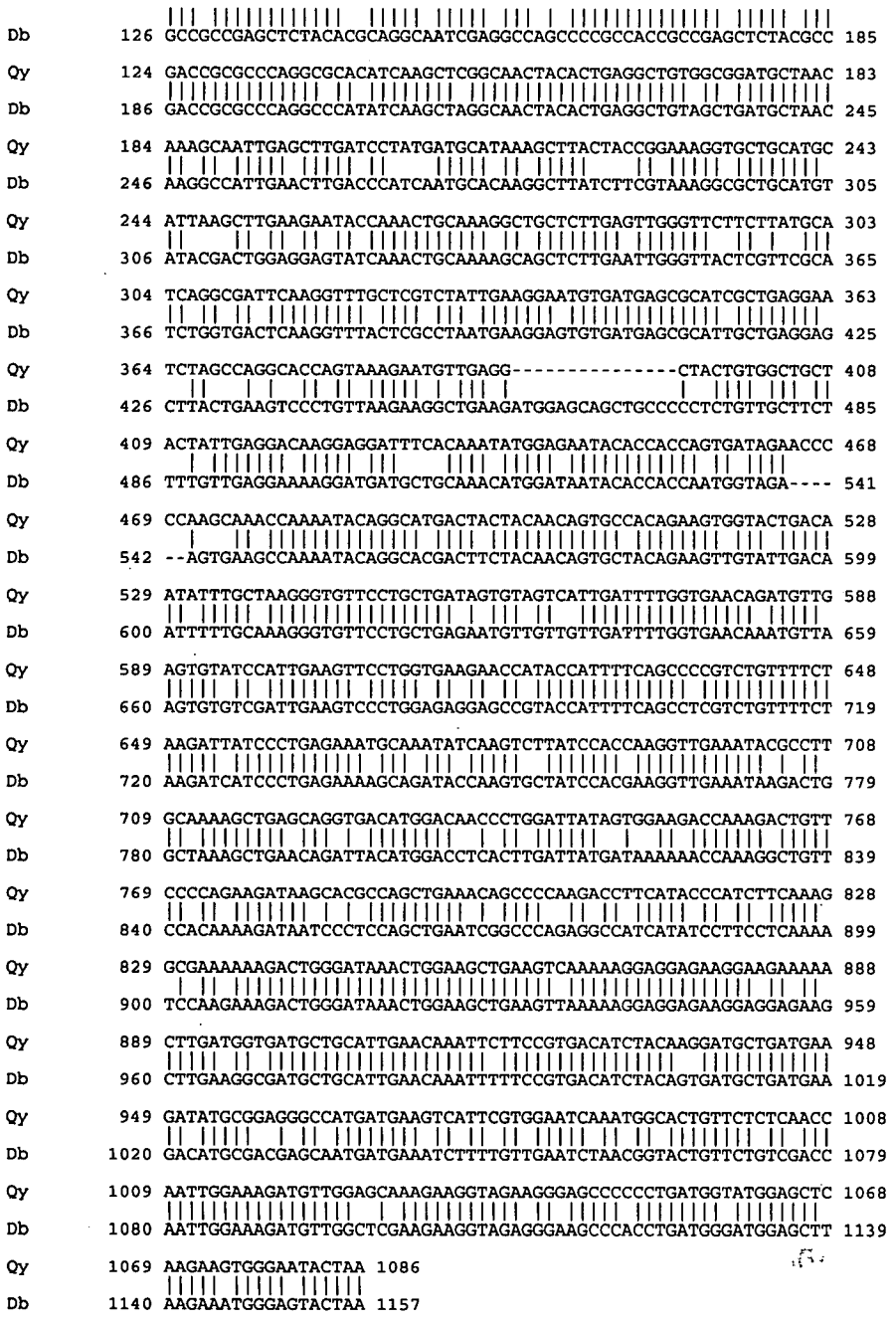
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	734.2	67.6	1409	8	AF439974	AF439974 Hordeum v
2	733.6	67.6	1222	8	AF192467	AF192467 Oryza sat
3	456.2	42.0	984	8	RID251317	AJ251317 Rubus ida
4	454.8	41.9	1113	8	AF494083	AF494083 Nicotiana
5	419.2	38.6	1366	8	AY085464	AY085464 Arabidops
6	418.2	38.5	1354	8	ATH318019	AJ318019 Arabidops
7	417.6	38.5	1077	6	AX412470	AX412470 Sequence
8	417.6	38.5	1108	8	AY150487	AY150487 Arabidops
9	417.6	38.5	1290	8	AF439976	AF439976 Arabidops
10	417.6	38.5	1400	8	AF428340	AF428340 Arabidops
11	417.6	38.5	1420	8	AF370229	AF370229 Arabidops
12	413.8	38.1	1197	8	AJ620883	AJ620883 Brassica
13	407.2	37.5	1068	8	AJ620882	AJ620882 Brassica
14	394.8	36.4	1291	8	AF439975	AF439975 Arabidops
15	393.2	36.2	1084	8	AY096395	AY096395 Arabidops
16	393.2	36.2	1348	8	AY063978	AY063978 Arabidops
17	285.8	26.3	968	8	ATH318020	AJ318020 Arabidops
18	260.4	24.0	582	8	AK059268	AK059268 Oryza sat
19	207.8	19.1	634	8	AF516180	AF516180 Nicotiana
20	207.8	19.1	634	8	AF516181	AF516181 Nicotiana
21	164.4	15.1	272	6	AR249214	AR249214 Sequence
22	159.8	14.7	1606	5	AJ720706	AJ720706 Gallus ga
23	156.4	14.4	1002	6	CQ724202	CQ724202 Sequence
24	156.4	14.4	1002	9	AF132856	AF132856 Homo sapi
25	156.4	14.4	1025	6	BD209970	BD209970 Human tra
26	156.4	14.4	1235	9	AY321358	AY321358 Homo sapi
27	156.4	14.4	1535	9	BC000911	BC000911 Homo sapi
28	156.4	14.4	1555	9	HSA344097	AJ344097 Homo sapi
29	156.4	14.4	1776	6	CQ490463	CQ490463 Sequence
30	156.4	14.4	1776	6	CQ496307	CQ496307 Sequence
31	156.4	14.4	2053	6	BD192630	BD192630 123 human
32	156.4	14.4	2053	6	AR184101	AR184101 Sequence
33	156.2	14.4	1002	9	BT009798	BT009798 Homo sapi
34	152.6	14.1	1518	5	BC072118	BC072118 Xenopus l
35	146.6	13.5	139201	8	AP002970	AP002970 Oryza sat
36	146.6	13.5	150022	8	AP003213	AP003213 Oryza sat
37	145.2	13.4	1314	5	CR761938	CR761938 Xenopus t
38	139.4	12.8	1104	10	BC009167	BC009167 Mus muscu
39	137.6	12.7	153865	8	AP002744	AP002744 Oryza sat
40	119	11.0	1089	9	AF068289	AF068289 Homo sapi
41	114.6	10.6	154558	8	AC107206	AC107206 Oryza sat
c 42	111	10.2	137843	9	AC138972	AC138972 Homo sapi
43	107.8	9.9	304	4	BTA344098	AJ344098 Bos tauru
44	90.8	8.4	1242	6	AX488766	AX488766 Sequence
45	89	8.2	110000	8	CR382128_09	Continuation (10 o

ALIGNMENTS

RESULT 1
AF439974
LOCUS AF439974 1409 bp mRNA linear PLN 06-NOV-2003
DEFINITION Hordeum vulgare SGT1 (SGT1) mRNA, complete cds.
ACCESSION AF439974
VERSION AF439974.1 GI:17017305
KEYWORDS .
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1409)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance
JOURNAL Science 295 (5562), 2073-2076 (2002)
MEDLINE 21893744
PUBMED 11847307
REFERENCE 2 (bases 1 to 1409)
AUTHORS Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A., Shirasu,K. and Schulze-Lefert,P.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich NR47UH, UK
FEATURES Location/Qualifiers
source 1. .1409
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
gene 1. .1409
/gene="SGT1"
CDS 9. .1130
/gene="SGT1"
/codon_start=1
/product="SGT1"
/protein_id="AAL33610.1"
/db_xref="GI:17017306"
/translation="MAAAAASDLESKAKEAPVDDDFELAAELYTQAIEAGPATAELYA
DRAQAHIKLGSYTEAVADANKAIELDPSMHKAYLRKGSACIKLEEYQTAKAALEVGSS
YASGDSRFTRLMKECDDRIAEEASQAPVKNAAAAVAPATSSGATTVVTEAEDQDGENM

```
                   ENAQPTVEVPSKPKYRHDYYNTPTEVVLTIFAKGVPADSVVVDFGEQMLSVSIELPGE
                   EPYHFQPRLFSKIVPDKCKYTVLSTKVEIRLAKAEPVTWTSLDYTGKPKAPQKINVPA
                   ESAQRPSYPSSKSKKDWDKLEAEVKKQEKDEKLDGDAALNKFFREIYSDADEDMRRAM
                   MKSFVESNGTVLSTNWKDVGKKTVEGSPPDGMELKKWEY"
ORIGIN

  Query Match             67.6%;  Score 734.2;  DB 8;  Length 1409;
  Best Local Similarity   80.6%;  Pred. No. 1.1e-121;
  Matches  897;  Conservative    0;  Mismatches  183;  Indels   33;  Gaps    2;

Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTA 80

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||| |||||||||||| || || || |||||||| || ||||| |||
Db         81 GCCGCCGAGCTCTACACCCAGGCCATCGAGGCTGGCCCAGCCACCGCGGAGCTCTACGCC 140

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              ||||| |||||||| |||||||||||| ||||  |||||||||||||| || ||||| |||
Db        141 GACCGAGCCCAGGCTCACATCAAGCTGGGCAGTTACACTGAGGCTGTAGCTGATGCCAAC 200

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              |||||||||||| ||||||||   |||||| ||||| |||   ||||| ||| |||| |||
Db        201 AAAGCAATTGAACTTGATCCCTCGATGCACAAAGCGTACCTTCGGAAGGGTTCTGCTTGC 260

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              || ||||| || |||||||||||||||||||||||||||||  | |||||||| ||||||
Db        261 ATCAAGCTGGAGGAATACCAAACTGCAAAGGCTGCTCTTGAAGTTGGTTCTTCCTATGCA 320

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              || || || ||||||||| |||||||  |||||||  |||||||| || || |||||||||
Db        321 TCTGGTGACTCAAGGTTTACTCGTCTTATGAAGGAGTGTGATGATCGTATTGCTGAGGAG 380

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCT------------------ 405
              |||||||||| ||||||||||||  |  ||  |||| |||
Db        381 GCTAGCCAGGCGCCAGTAAAGAATGCCGCTGCGGCTGTTGCTCCTGCTACATCTTCCGGG 440

Qy        406 ------------GCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCA 453
                          | ||||   |  ||  |  |||||     |||||||||||||| |||
Db        441 GCTACAACTGTGGTTACTGAAGCTGAGGACCAGGATGGTGAAAATATGGAGAATGCACAG 500

Qy        454 CCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACA 513
              |||  | |||||   ||||||||| || ||||||||||||||||||||||| | | |||
Db        501 CCAACGGTAGAAGTGCCAAGCAAGCCCAAATACAGGCATGACTACTACAATACTCCTACA 560

Qy        514 GAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTT 573
              |||||||||||||| ||||||||||||||||||||| ||||| || || ||  |||| |||
Db        561 GAAGTGGTACTGACTATATTTGCTAAGGGTGTTCCAGCTGACAGCGTGGTTGTTGACTTT 620

Qy        574 GGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAG 633
              |||||||||||| ||||||| || |||||| |||||||||| ||||||||||||| |||
Db        621 GGTGAACAGATGCTGAGTGTCTCAATTGAACTTCCTGGTGAGGAACCATACCATTTCCAG 680

Qy        634 CCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAG 693
              || |||||||||| |||||  |||| || || ||||| |||   || || || || |||
Db        681 CCTCGTCTGTTTTCAAAGATCGTCCCAGATAAGTGCAAGTATACTGTATTGTCTACAAAG 740

Qy        694 GTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGA 753
              || |||||||||||.|||||||||||||   || || |||||| | |||||||||| |||
Db        741 GTCGAAATACGCCTGGCAAAAGCTGAGCCAGTAACTTGGACATCACTGGATTATACTGGT 800

Qy        754 AGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCA 813
              | |||||||   | ||| ||||||||||     |||||||| ||||||  || || ||
Db        801 AAACCAAAG---GCTCCTCAGAAGATAAATGTACCAGCTGAATCAGCCCAGAGGCCATCT 857

Qy        814 TACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAG 873
              || || || |||||  | ||||| |||||||||| || || |||||||| |||||  ||
Db        858 TATCCTTCATCAAAATCCAAAAAGGACTGGGATAAGCTTGAGGCTGAAGTGAAAAAACAG 917

Qy        874 GAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTAC 933
              |||||||| || |||||||| ||||||||||||||||||| || |||||||| || |||
Db        918 GAGAAGGATGAGAAACTTGACGGTGATGCTGCATTGAACAAGTTTTTCCGTGAAATTTAC 977

Qy        934 AAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGC 993
              |  |||||||||||||||||||| || || |||||||||| || ||||| || |||||
Db        978 AGTGATGCTGATGAAGATATGCGTAGAGCAATGATGAAGTCTTTTGTGGAGTCTAATGGA 1037

Qy        994 ACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCT 1053
              || |||||||||||||| |||||||||| ||   ||||| ||| ||||| ||||| |||
Db       1038 ACCGTTCTCTCAACCAACTGGAAAGATGTCGGGAAAAAGACGGTTGAAGGAAGCCCTCCT 1097

Qy       1054 GATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              ||||| |||||||||||||||||||| || |||
Db       1098 GATGGAATGGAGCTCAAGAAGTGGGAGTATTAA 1130


RESULT 2
AF192467
LOCUS       AF192467                1222 bp    mRNA    linear   PLN 16-DEC-1999
DEFINITION  Oryza sativa Sgt1 (Sgt1) mRNA, complete cds.
ACCESSION   AF192467
VERSION     AF192467.1  GI:6581057
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1222)
  AUTHORS   Shirasu,K., Lahaye,T., Tan,M.W., Zhou,F., Azevedo,C. and
            Schulze-Lefert,P.
  TITLE     A novel class of eukaryotic zinc-binding proteins is required for
            disease resistance signaling in barley and development in C.
            elegans
  JOURNAL   Cell 99 (4), 355-366 (1999)
  MEDLINE   20035956
   PUBMED   10571178
REFERENCE   2  (bases 1 to 1222)
  AUTHORS   Shirasu,K. and Schulze-Lefert,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-1999) The Sainsbury Laboratory, John Innes
            Center, Colney Lane, Norwich NR4 7UH, United Kingdom
FEATURES             Location/Qualifiers
     source          1. .1222
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:4530"
                     /note="corresponds to MAFF DNA Bank EST clone C62803"
     gene            1. .1222
                     /gene="Sgt1"
     CDS             54. .1157
                     /gene="Sgt1"
                     /note="similar to yeast Sgt1"
                     /codon_start=1
                     /product="Sgt1"
                     /protein_id="AAF18438.1"
                     /db_xref="GI:6581058"
                     /translation="MATAAASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYA
                     DRAQAHIKLGNYTEAVADANKAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYS
                     FASGDSRFTRLMKECDERIAEELTEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPP
                     MVEVKPKYRHDFYNSATEVVLTIFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQP
                     RLFSKIIPEKSRYQVLSTKVEIRLAKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRP
                     SYPSSKSKKDWDKLEAEVKKEEKEEKLEGDAALNKFFRDIYSDADEDMRRAMMKSFVE
                     SNGTVLSTNWKDVGSKKVEGSPPDGMELKKWEY"
ORIGIN

  Query Match             67.6%;  Score 733.6;  DB 8;  Length 1222;
  Best Local Similarity   80.9%;  Pred. No. 1.5e-121;
  Matches  888;  Conservative    0;  Mismatches  189;  Indels   21;  Gaps    2;

Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         66 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGACGACTTCGAGCTC 125

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
```



RESULT 3

RID251317
LOCUS RID251317 984 bp mRNA linear PLN 24-NOV-1999
DEFINITION Rubus idaeus partial mRNA for putative protein phosphatase.
ACCESSION AJ251317
VERSION AJ251317.1 GI:6468694
KEYWORDS protein phosphatase.
SOURCE Rubus idaeus
ORGANISM Rubus idaeus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rubus.
REFERENCE 1
AUTHORS Schroeder,G., Lurz,G. and Schroeder,J.
TITLE cDNA for protein phosphatase like protein from raspberry (Rubus idaeus)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 984)
AUTHORS Schroeder,G.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1999) Schroeder G., Institut fuer Biologie II, Universitaet Freiburg, Schaenzlestr. 1, Freiburg D-79104, Germany
FEATURES Location/Qualifiers
source 1. .984
/organism="Rubus idaeus"
/mol_type="mRNA"
/db_xref="taxon:32247"
CDS <1. .984
/codon_start=1
/product="putative protein phosphatase"
/protein_id="CAB61630.1"
/db_xref="GI:6468695"
/db_xref="UniProt/TrEMBL:Q9SM53"
/translation="PKNSELYSDRAQANMKSNNLTEAVADANKAIELDPSLFKAYLRK
GIACMKLEEYRTAKAALELGASLAPGESRFAKMIKECDELIAEENGEPPKQPMETTTT
EIVAEDVEPVDPPSNEVTVAPVKPKYRHEFYQKAEEVVVTIFAKGIPAENVAVDFGPQ
ILSVSIDVPGEVHIHFQPRLFGKIIPEKCRFEVLSTKVEIRLAKAEPNHWTSLEFSKD
NPVPLKVNAPVIGAQRPSYPSSKPKRVDWDKLEAQVKKEEKDEKLDGDAALNKFFQDI
YKDADEDTRRAMRKSFVESNGTVLSTNWKEVGNKKVEGSAPDGMEMKKWEI"
ORIGIN

Query Match 42.0%; Score 456.2; DB 8; Length 984;
Best Local Similarity 68.5%; Pred. No. 7.9e-72;
Matches 677; Conservative 0; Mismatches 303; Indels 9; Gaps 3;

```
Db        418 GTTGTGACTATATTTGCAAAGGGCATACCTGCCGAAAATGTTGCTGTTGATTTTGGTCCA 477

Qy        580 CAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGT 639

Db        478 CAAATATTAAGTGTCAGCATTGATGTTCCCGGTGAAGTACATATACATTTTCAGCCTCGC 537

Qy        640 CTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAA 699

Db        538 CTATTTGGAAAGATCATTCCTGAAAAGTGCAGATTTGAGGTTTTGTCTACCAAAGTTGAA 597

Qy        700 ATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCA 759

Db        598 ATTCGCCTTGCAAAAGCTGAACCTAATCACTGGACTTCTCTTGAGTTCA---GCAAGGAT 654

Qy        760 AAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCA 819

Db        655 AACCCTGTTCCTCTAAAGGTTAATGCCCCAGTTATTGGAGCCCAAAGGCCATCTTACCCA 714

Qy        820 TCTTCAAAGGCGAAAA---AAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 876

Db        715 TCCTCCAAACCAAAAAGGGTAGATTGGGACAAGCTTGAAGCCCAAGTGAAGAAGGAGGAA 774

Qy        877 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 936

Db        775 AAAGATGAGAAGCTTGATGGTGATGCAGCTTTGAACAAATTTTTCCAAGATATATACAAG 834

Qy        937 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 996

Db        835 GATGCTGATGAAGACACAAGAAGGGCCATGAGAAAATCTTTTGTGGAGTCAAATGGGACG 894

Qy        997 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGAT 1056

Db        895 GTGCTTTCAACAAACTGGAAAGAAGTGGGCAACAAGAAGGTGGAGGGAAGTGCCCCTGAT 954

Qy       1057 GGTATGGAGCTCAAGAAGTGGGAATACTA 1085

Db        955 GGTATGGAGATGAAGAAATGGGAGATCTA 983
```

```
RESULT 4
AF494083
LOCUS       AF494083               1113 bp    mRNA    linear   PLN 30-OCT-2003
DEFINITION  Nicotiana benthamiana SGT1 mRNA, complete cds.
ACCESSION   AF494083
VERSION     AF494083.1  GI:29468338
KEYWORDS    .
SOURCE      Nicotiana benthamiana
  ORGANISM  Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1113)
  AUTHORS   Liu,Y., Schiff,M., Serino,G., Deng,X.W. and Dinesh-Kumar,S.P.
  TITLE     Role of SCF ubiquitin-ligase and the COP9 signalosome in the N
            gene-mediated resistance response to Tobacco mosaic virus
  JOURNAL   Plant Cell 14 (7), 1483-1496 (2002)
  MEDLINE   22115173
   PUBMED   12119369
REFERENCE   2  (bases 1 to 1113)
  AUTHORS   Liu,Y., Schiff,M., Serino,G., Dinesh-Kumar,S.P. and Deng,X.-W.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) MCDB, Yale University, New Haven, CT 06520,
            USA
FEATURES             Location/Qualifiers
     source          1. .1113
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
     CDS             1. .1113
                     /codon_start=1
                     /product="SGT1"
                     /protein_id="AAO85509.1"
                     /db_xref="GI:29468339"
                     /translation="MASDLETRAKEAFIDDHFELAVDLYTQAIAMTPKNAELFADRAQ
                     ANIKLNYFTEAVVDANKAIELDPSMSKAYLRKGLACMKLEEYQTAKAALETGASLAPA
                     ESRFTKLIKECDERIAEEAGELPNQSVDKTSGNVVAPPASESLDNVAVAPKDAQPTVN
                     LSYQGSAARPKYRHEFYQKPEEVVVTIFAKGIPAKNVVVDFGEQILSVSIDVPGDETY
                     SFQPRLFGKITPAKCRYEVMSTKIEIRLAKAEPLHWTSLEYTRASAVVQRPNVSSDAP
                     RPSYPSSKLRHVDWDKLEAEVKKEEKDEKLDGDAALNKFFRDIYKDADEDTRRAMMKS
                     FVESNGTVLSTNWKEVGTKKVEGSPPDGMELKKWEI"
ORIGIN
```

```
  Query Match             41.9%;  Score 454.8;  DB 8;  Length 1113;
  Best Local Similarity   65.6%;  Pred. No. 1.4e-71;
  Matches  729;  Conservative    0;  Mismatches  347;  Indels   36;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66

Db          4 GCGTCCGATCTGGAGACTAGGGCTAAAGAAGCTTTCATTGACGACCACTTTGAGCTCGCC 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126

Db         64 GTTGACCTTTACACTCAAGCAATTGCCATGACTCCTAAGAACGCTGAGCTTTTCGCCGAC 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186

Db        124 CGTGCTCAGGCCAACATCAAACTCAACTACTTCACTGAAGCTGTTGTTGATGCGAACAAG 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246

Db        184 GCCATTGAGTTAGATCCTTCAATGTCAAAAGCATATTTGCGTAAGGGGTTGGCCTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306

Db        244 AAGCTTGAAGAGTACCAAACTGCAAAAGCAGCTTTGGAAACTGGTGCTAGTTTAGCACCG 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366

Db        304 GCAGAGTCAAGGTTCACAAAGTTAATCAAAGAATGTGATGAACGCATTGCAGAGGAAGCT 363

Qy        367 AGCCAGGCACCAGTAAAG----------------------AATGTTGAGGCTACTGTGGCT 405

Db        364 GGAGAACTACCTAATCAGTCGGTGGATAAAACCTCGGGAAATGTCGTAGCTCCCCCTGCA 423

Qy        406 GCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGA----- 460

Db        424 TCTGAGTCTTTGGACAATGTTGCTGTTGCGCCTAAAGATGCTCAACCAACTGTCAACCTG 483

Qy        461 -------TAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACA 513

Db        484 TCCTATCAAGGATCTGCTGCCAGACCAAAATACAGGCATGAATTTTACCAGAAGCCAGAG 543

Qy        514 GAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTT 573

Db        544 GAGGTGGTGGTGACTATATTTGCCAAGGGAATACCAGCCAAGAATGTTGTTGTTGACTTT 603

Qy        574 GGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAG 633

Db        604 GGTGAACAAATACTTAGTGTTAGCATTGATGTGCCGGGTGACGAAACTTATTCTTTCCAG 663

Qy        634 CCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAG 693

Db        664 CCTAGGTTGTTTGGCAAGATAACACCTGCAAAATGCAGATATGAAGTGATGTCCACCAAA 723

Qy        694 GTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGA 753

Db        724 ATTGAGATCCGCCTTGCAAAAGCTGAACCCTTACACTGGACATCTCTTGAATATA---CG 780

Qy        754 AGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCA 813

Db        781 AGAGCGTCTGCTGTAGTGCAGAGGCCTAATGTGTCATCAGATGCCCCCCGGCCCAGTTAT 840

Qy        814 TACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAG 873

Db        841 CCTTCCTCGAAATTGAGACATGTGGATTGGGATAAATTAGAGGCTGAAGTGAAAAAGGAG 900
```

```
Qy        874 GAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTAC 933
              || || || |||||| | ||||| ||||| ||||||||||||| ||||| ||||| |||
Db        901 GAAAAAGACGAAAAATTGGATGGAGATGCAGCATTGAACAAATTTTTCCGAGACATTTAC 960

Qy        934 AAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGC 993
              || ||||| ||||| || |      | ||  ||||||||||| || || |||||||| |||||
Db        961 AAAGATGCCGATGAGGACACCAGAAGAGCCATGATGAAATCCTTTGTGGAATCTAATGGG 1020

Qy        994 ACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCT 1053
              ||||||  | || || || |||||||| |||||  ||||||||| ||||| ||||| ||
Db       1021 ACTGTTTTGTCGACAAACTGGAAAGAAGTTGGCACAAAGAAGGTCGAAGGAAGCCCTCCG 1080

Qy       1054 GATGGTATGGAGCTCAAGAAGTGGGAATACTA 1085
              ||||| |||||||| ||||| |||||    |||
Db       1081 GATGGCATGGAGCTGAAGAAATGGGAGATCTA 1112


RESULT 5
AY085464
LOCUS       AY085464                1366 bp    mRNA    linear   PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 152712 mRNA, complete sequence.
ACCESSION   AY085464
VERSION     AY085464.1  GI:21404174
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1366)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
  JOURNAL   Genome Biol. 3 (6), RESEARCH0029 (2002)
  MEDLINE   22088475
   PUBMED   12093376
REFERENCE   2  (bases 1 to 1366)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-Length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1366)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1..1366
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="152712"
     CDS             57..1133
                     /codon_start=1
                     /product="SGT1a"
                     /protein_id="AAM62690.1"
                     /db_xref="GI:21553597"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match             38.6%;  Score 419.2;  DB 8;  Length 1366;
  Best Local Similarity   63.1%;  Pred. No. 3.4e-65;
  Matches  681;  Conservative    0;  Mismatches  393;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              |||  |||  | |        || || || || || || ||   | || ||||||||  | ||
Db         60 GCGAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 119

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
                  ||  |  |||  | | ||||| |||    |  |||     ||||| | |||  |||||
Db        120 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 179

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              || || |||||  |||||||  |||  |||| |||||| |||||  |  ||||| ||||||
Db        180 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 239

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || |||||| | || ||  | | ||    ||||| ||    |  | | ||  |||| || ||
Db        240 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 299

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
              ||||| ||||||||    ||||| || ||  || ||  ||  |||  ||||| ||||  |
Db        300 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 359

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
                  ||  ||| ||||           |  |   | |||||  |||    || || || || |||
Db        360 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 419

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
              |    ||         ||  ||      |       |  |  | || ||  |  |         |
Db        420 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 476

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
              |        |      |      |     | |       ||| |       | | | ||    |||||| |   | |
Db        477 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 536

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
              ||||| ||  ||  | |||| | |       ||||| |||||| ||  ||         ||
Db        537 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 596

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
              |||    || |    |||      | ||  |||||||| |||||  |||||||     |||||| |||
Db        597 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 656

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
                ||||  ||  |||  |  || ||  |||  | |||||    |    ||||      ||||| ||  || |||||
Db        657 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 716

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
              |||| ||  |  ||||  ||  ||  ||||| |||||  ||  ||  ||||| ||||| |||        |
Db        717 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 776

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
              || |||  |   |||  |||  || ||   ||  | ||| |  |       | | | |  ||   |     |
Db        777 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 836

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
               ||||              | |   |||||       |||||||||| |||  |        |||||||||||
Db        837 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 893

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
```

```
Db        894 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 953

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966

Db        954 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1013

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026

Db       1014 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1073

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086

Db       1074 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1133


RESULT 6
ATH318019
LOCUS       ATH318019              1354 bp    mRNA    linear   PLN 23-OCT-2003
DEFINITION  Arabidopsis thaliana mRNA for SGT1-like protein (sgt1a gene).
ACCESSION   AJ318019
VERSION     AJ318019.1  GI:30524963
KEYWORDS    SGT1-like protein; sgt1a gene.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Pawlowski,J., Holzmann,M., Berney,C., Fahrni,J., Gooday,A.J.,
            Cedhagen,T., Habura,A. and Bowser,S.S.
  TITLE     The evolution of early Foraminifera
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11494-11498 (2003)
  MEDLINE   22882917
   PUBMED   14504394
REFERENCE   2  (bases 1 to 1354)
  AUTHORS   Shen,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) Shen W.H., Biologie Cellulaire, IBMP/CNRS,
            12, rue du General Zimmer, 67084 Strasbourg, FRANCE
FEATURES             Location/Qualifiers
     source          1. .1354
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1. .1354
                     /gene="sgt1a"
     CDS             40. .1116
                     /gene="sgt1a"
                     /function="proteolysis"
                     /codon_start=1
                     /product="SGT1-like protein"
                     /protein_id="CAC85266.1"
                     /db_xref="GI:30524964"
                     /db_xref="UniProt/TrEMBL:Q9SUT5"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"

ORIGIN

  Query Match           38.5%;  Score 418.2;  DB 8;  Length 1354;
  Best Local Similarity  63.0%;  Pred. No. 5.1e-65;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66

Db         43 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 102

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126

Db        103 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 162

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186

Db        163 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 222

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246

Db        223 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 282

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306

Db        283 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 342

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366

Db        343 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 402

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426

Db        403 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 459

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486

Db        460 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 519

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546

Db        520 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 579

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606

Db        580 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 639

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666

Db        640 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 699

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726

Db        700 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 759

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786

Db        760 ACNTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 819

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846

Db        820 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 876

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906

Db        877 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 936

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966

Db        937 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 996

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026

Db        997 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1056

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086

Db       1057 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1116


RESULT 7
AX412470
LOCUS       AX412470               1077 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 234 from Patent WO0222675.
ACCESSION   AX412470
```

```
VERSION     AX412470.1  GI:21444928
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
  TITLE     Plant genes, the expression of which are altered by pathogen
            infection
  JOURNAL   Patent: WO 0222675-A 234 21-MAR-2002;
            Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
            CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
            Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES             Location/Qualifiers
     source          1. .1077
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN

  Query Match           38.5%;  Score 417.6;  DB 6;  Length 1077;
  Best Local Similarity  63.0%;  Pred. No. 6.5e-65;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   |||  | |      || || || || || ||  | || || ||||||||  | ||
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
                ||   | |||  | | ||||| |||    |  |||      |||||  | |||  |||||
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              || || |||||  |||||||  |||  |||| |||||| |||||  |  ||||| ||||||
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || |||||| | || || | | ||   ||||| ||   | | || ||   |||| || ||
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
              ||||| ||||||||     |||||  || || || || ||    |||  |||||  |||| |
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
              ||  |||  ||||      |  |  | ||||| |||   || || || || |||
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
              |   | |      ||  ||      | | | || || | |     |
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
              |       | |   |    | |     ||| |    | | | ||   |||||| |  | |
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
              |||||  ||  | ||||  | |    |   |||||  |||  |||||||  ||  ||  ||    ||
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
              |||   || |  |||      | ||  |||||||| |||||   |||||||    ||||||  ||||
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
              ||||  ||  |||  | ||  |||  | |||||   |  ||||     |||||  ||  || |||||
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
              ||||  ||  | ||||  ||  ||  |||||  |||||  ||  ||  |||||  |||||  |||     |
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
              ||  |||  |  ||||  ||  |||  ||  |  |  |  ||     |  |||  |      |  |
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
              ||||        |  |  |||||    ||||||||||| |||  |   ||||||||||||
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
              ||  |||||||||||||| ||  ||  |||||||||| || || ||||||||  ||||| ||
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
              ||||||||||  |||  |  |||||  |||   |  |||  ||||||||  |||  ||  ||||  |||
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
              |  ||  |||||  |  ||  ||  |||||  ||  ||  ||  ||  ||  ||  |||||||||  |||||
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              | |||||  ||  ||  |  |  ||  ||  |||||  ||||||||||||||||||||  ||  |  |
Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077
```

```
RESULT 8
AY150487
LOCUS       AY150487               1108 bp    mRNA    linear   PLN 23-SEP-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete
            cds.
ACCESSION   AY150487
VERSION     AY150487.1  GI:23297701
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C., Toriumi,M., Wallender,B.K., Wong,C., Wu,H.C.,
            Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
            Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
            Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1108)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J.,
            Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
            Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
```

```
            Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
            to this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
            /PGEC) contributed equally to this work as PIs.

            Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1..1108
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="U09567"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1..1108
                     /gene="At4g11260"
     CDS             1..1077
                     /gene="At4g11260"
                     /codon_start=1
                     /evidence=experimental
                     /product="unknown protein"
                     /protein_id="AAN12904.1"
                     /db_xref="GI:23297702"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
     3'UTR           1078..1108
                     /gene="At4g11260"
ORIGIN

  Query Match            38.5%;  Score 417.6;  DB 8;  Length 1108;
  Best Local Similarity  63.0%;  Pred. No. 6.5e-65;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077


RESULT 9
AF439976
LOCUS       AF439976                1290 bp    mRNA    linear   PLN 06-NOV-2003
DEFINITION  Arabidopsis thaliana SGT1b (SGT1b) mRNA, complete cds.
ACCESSION   AF439976
VERSION     AF439976.1  GI:17017309
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1290)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     The RAR1 interactor SGT1, an essential component of R
            gene-triggered disease resistance
  JOURNAL   Science 295 (5562), 2073-2076 (2002)
  MEDLINE   21893744
   PUBMED   11847307
REFERENCE   2  (bases 1 to 1290)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich NR47UH, UK
FEATURES             Location/Qualifiers
     source          1..1290
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1..1290
                     /gene="SGT1b"
```

```
CDS             1. .1077
                /gene="SGT1b"
                /codon_start=1
                /product="SGT1b"
                /protein_id="AAL33612.1"
                /db_xref="GI:17017310"
                /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                WKEVGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match          38.5%;  Score 417.6;  DB 8;  Length 1290;
  Best Local Similarity   63.0%;  Pred. No. 6.5e-65;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db          4 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db         64 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db        124 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db        184 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db        244 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db        304 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 363

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db        364 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 420

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db        421 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 480

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db        481 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 540

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db        541 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 600

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db        601 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 660

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db        661 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 720

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db        721 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 780

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db        781 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 837

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db        838 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 897

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db        898 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 957

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db        958 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1017

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db       1018 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1077


RESULT 10
AF428340
LOCUS       AF428340                1400 bp    mRNA    linear   PLN 18-OCT-2001
DEFINITION  Arabidopsis thaliana AT4g11260/F8L21_50 mRNA, complete cds.
ACCESSION   AF428340
VERSION     AF428340.1  GI:16226817
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1400)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1400)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
            Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J. Bowser,L.,
            Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S.,
            Davis,R.W.,  Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
```

```
          contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1400
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL09-50-H05(R14746)"
                     /ecotype="Columbia"
     5'UTR           1. .131
     CDS             132. .1208
                     /note="unknown protein"
                     /codon_start=1
                     /product="AT4g11260/F8L21_50"
                     /protein_id="AAL16270.1"
                     /db_xref="GI:16226818"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
     misc_difference 703
                     /note="compared to genome"
                     /replace="g"
     3'UTR           1209. .1400
ORIGIN

  Query Match            38.5%;  Score 417.6;  DB 8;  Length 1400;
  Best Local Similarity  63.0%;  Pred. No. 6.5e-65;
  Matches  680;  Conservative    0;  Mismatches  394;  Indels    6;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              ||   |||  | |     || || || || || ||  | || || ||||||||  | ||
Db        135 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 194

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
                ||  | |||  | | ||||| |||   |  |||     |||||  | |||  |||||
Db        195 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 254

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              || || |||||  |||||||  |||  |||| |||||| ||||| |  ||||| ||||||
Db        255 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 314

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              || |||||| | || || | | ||   ||||| ||   | | || ||  |||| || ||
Db        315 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 374

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
              ||||| ||||||||    |||||  || || || || ||   |||  |||||  |||| |
Db        375 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 434

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
                ||  ||| ||||      |  |  | ||||| |||   || || || || |||
Db        435 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 494

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
              |  | |     ||  ||      | | | || ||  |     |
Db        495 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 551

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
              |       |  |    | |    ||| |    | | | ||   |||||| |  | |
Db        552 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 611

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
              ||||| || | |||| | |    |  |||| ||| ||||||| || || ||    ||
Db        612 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 671

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
              |||   || |  |||    | || |||||||| |||||  |||||||   |||||| |||
Db        672 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 731

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
              |||| || ||| | || ||| | |||||  |  ||||    ||||| || || |||||
Db        732 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 791

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
              |||| || | |||| || || ||||| ||||| || || ||||| ||||| |||    |
Db        792 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 851

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
              || ||| |  |||| || ||| ||  | | | ||    | | ||| |    |  | |
Db        852 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 911

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
              ||||      | |  |||||    |||||||||| ||| |   ||||||||||
Db        912 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 968

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
              ||  |||||||||||||| || ||  |||||||||| || || |||||||| ||||| ||
Db        969 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 1028

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
              |||||||||| ||| | ||||| |||   |  ||| |||||||| ||| || |||| |||
Db       1029 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1088

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
              |  || ||||| |  || || ||||| || || || || || || |||||||| |||||
Db       1089 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1148

Qy        1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
               | ||||| || ||  | |  || || ||||| ||||||||||||||||||||| || | |
Db        1149 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1208

RESULT 11
AF370229
LOCUS       AF370229               1420 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g11260) mRNA, complete
            cds.
ACCESSION   AF370229
VERSION     AF370229.1  GI:13877932
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1420)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1420)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2001) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
```

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

```
FEATURES             Location/Qualifiers
     source          1. .1420
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL04-14-O07 (R09567)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (Lambda ZAP) as a XhoI/SstI insert."
     gene            1. .1420
                     /gene="At4g11260"
     5'UTR           1. .136
                     /gene="At4g11260"
     misc_difference 32. .34
                     /gene="At4g11260"
                     /note="not present in genomic sequence"
     CDS             137. .1213
                     /gene="At4g11260"
                     /codon_start=1
                     /evidence=experimental
                     /product="unknown protein"
                     /protein_id="AAK44044.1"
                     /db_xref="GI:13877933"
                     /translation="MAKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQ
                     ANIKIDNFTEAVVDANKAIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPN
                     EPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEADAPPVPIPAAPAKPMF
                     RHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIP
                     EKCRFEVLSTKVEIRLAKAEIITWASLEYGKGQSVLPKPNVSSALSQRPVYPSSKPAK
                     DWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTN
                     WKEVGTKKVESTPPDGMELKKWEY"
     3'UTR           1214. .1420
                     /gene="At4g11260"
ORIGIN
```

Query Match 38.5%; Score 417.6; DB 8; Length 1420;
Best Local Similarity 63.0%; Pred. No. 6.5e-65;
Matches 680; Conservative 0; Mismatches 394; Indels 6; Gaps 2;

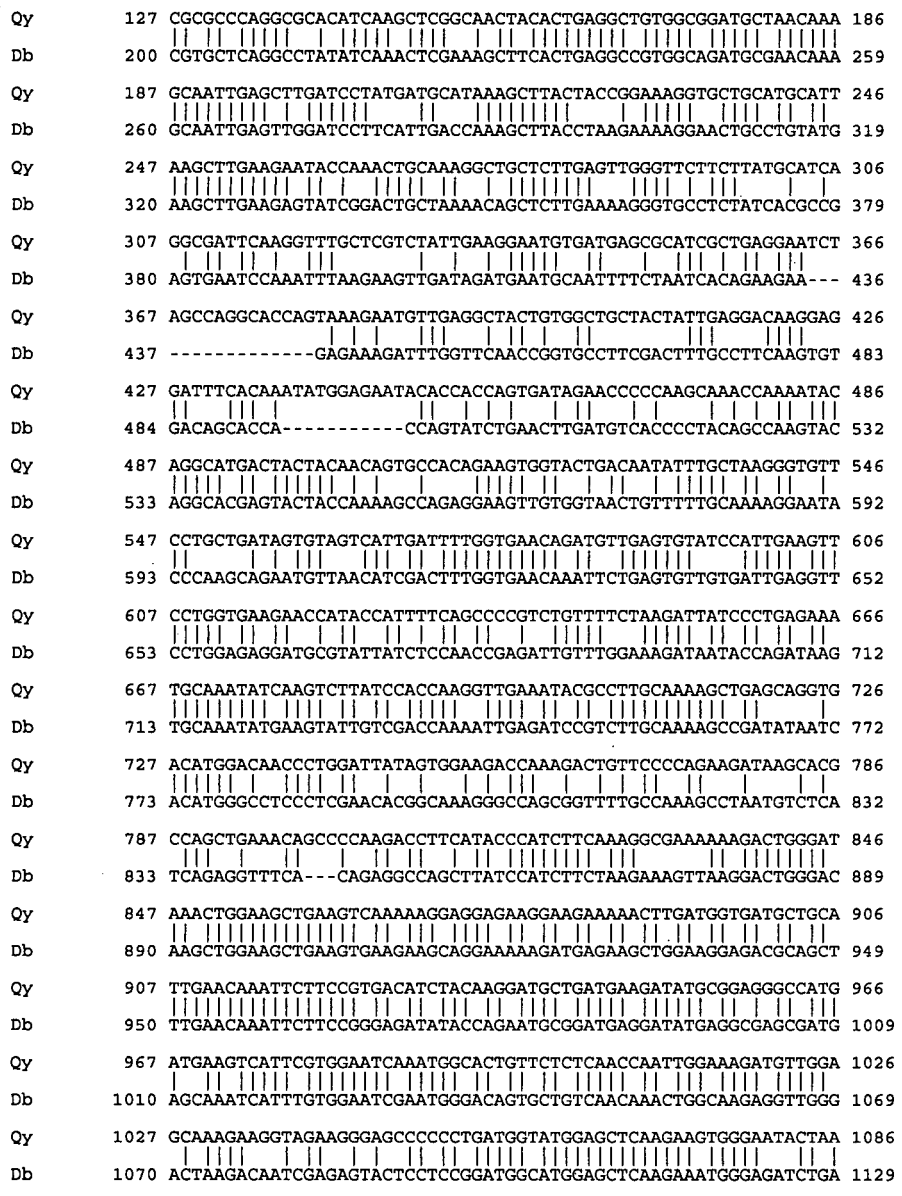
```
Qy        7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db      140 GCCAAGGAATTAGCAGAGAAAGCTAAAGAAGCTTTTCTAGATGATGACTTCGATGTTGCT 199

Qy       67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db      200 GTTGACTTATACTCCAAAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 259

Qy      127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db      260 CGTGCTCAGGCCAACATCAAAATCGATAACTTCACTGAAGCTGTTGTAGATGCGAACAAA 319

Qy      187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db      320 GCCATTGAGTTGGAGCCAACGTTGGCAAAAGCCTATCTCAGAAAGGGCACTGCTTGTATG 379

Qy      247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db      380 AAGCTAGAAGAATATAGTACTGCTAAAGCAGCCCTGGAAAAGGGAGCTTCTGTTGCACCG 439
```

```
Qy      307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db      440 AATGAACCAAAGTTTAAGAAGATGATAGATGAATGCGATCTTCGTATTGCAGAAGAAGAG 499

Qy      367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db      500 A---AAGATTTGGTTCAGCCGATGCCACCGAGTTTGCCTTCAAGCTCTACAACACCACTA 556

Qy      427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db      557 GCAACGGAAGCTGATGCTCCTCCTGTTCCAATTCCTGCAGCACCTGCCAAACCGATGTTC 616

Qy      487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db      617 AGGCACGAGTTCTACCAGAAACCAGAAGAAGCGGTGGTGACAATTTTCGCCAAAAAAGTA 676

Qy      547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db      677 CCTAAGGAGAACGTAACTGTCGAGTTTGGTGAGCAGATTCTGAGTGTTGTCATTGATGTT 736

Qy      607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db      737 GCTGGAGAGGAAGCTTATCATCTCCAGCCGAGATTGTTCGGGAAGATAATACCAGAGAAG 796

Qy      667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db      797 TGCAGATTTGAAGTATTGTCGACCAAAGTTGAGATCCGTCTTGCGAAAGCAGAGATAATC 856

Qy      727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db      857 ACCTGGGCCTCCCTTGAATATGGTAAAGGGCAAAGTGTTTTGCCCAAACCCAATGTCTCA 916

Qy      787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db      917 TCAGC---GCTGTCGCAGAGACCAGTGTACCCATCTTCTAAGCCAGCAAAAGACTGGGAC 973

Qy      847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db      974 AAGTTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTTGATGGAGATGCAGCT 1033

Qy      907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db     1034 ATGAACAAATTTTTCAGCGACATATACTCGAGTGCAGATGAAGACATGAGGCGGGCAATG 1093

Qy      967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db     1094 AACAAATCATTTGCAGAGTCGAATGGGACGGTACTGTCGACAAACTGGAAAGAAGTTGGG 1153

Qy     1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db     1154 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTCAAGAAGTGGGAGTATTGA 1213
```

```
RESULT 12
AJ620883
LOCUS       AJ620883               1197 bp    mRNA    linear   PLN 16-JAN-2004
DEFINITION  Brassica oleracea mRNA for SGT1-like protein (sgt1b-2 gene).
ACCESSION   AJ620883
VERSION     AJ620883.1  GI:40974916
KEYWORDS    SGT1-like protein; sgt1b-2 gene.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Yemm,A.A., Holub,E. and Tor,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1197)
  AUTHORS   Yemm,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,
            Horticultural Research International, Wellesbourne, Warwick, UNITED
```

```
                KINGDOM
FEATURES             Location/Qualifiers
     source          1..1197
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /variety="Senna"
                     /db_xref="taxon:3712"
                     /tissue_type="leaf"
     gene            1..1197
                     /gene="sgt1b-2"
     CDS             1..1065
                     /gene="sgt1b-2"
                     /function="role in disease resistance"
                     /codon_start=1
                     /evidence=experimental
                     /product="SGT1-like protein"
                     /protein_id="CAF06581.1"
                     /db_xref="GI:40974917"
                     /translation="MASELAEKAKEAFLDDDFDVAADLYSKAIDLDPSCASFFADRAQ
                     ANIKLLNFTEAVADANKAIELEPTLAKAYLRKGAACMKLEEYATAKAALEKGASVAPN
                     ESKFEKMIDECNLLIAEEEKDLVQQVPPTLPSSSTTPLAIAADAPPAAPAKPMFRHEF
                     YQKPEEVGVAIFAKGIPKQNVNVEFGDQILSVVIDVAGEEAYHFQPRLFGKIIPEKCR
                     YEVLSTKVEIRLAKAEIVTWASLEYGKGQALLPKPNVASAVSQRPVYPSSKPGKDWDK
                     LEAEVKKQEKDEKLDGDAAMNKFFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWKEV
                     GTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match             38.1%;  Score 413.8;  DB 8;  Length 1197;
  Best Local Similarity   63.6%;  Pred. No. 3.1e-64;
  Matches  687;  Conservative    1;  Mismatches  373;  Indels   20;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66

Db          4 GCGAGCGAACTAGCGGAAAAAGCTAAAGAGGCGTTCTTAGATGACGACTTCGATGTGGCT 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126

Db         64 GCTGACTTATACTCCAAAGCCATTGACTTGGATCCCTCTTGCGCCTCCTTCTTCGCCGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186

Db        124 CGTGCTCAAGCCAACATCAAACTCCTTAACTTCACCGAAGCTGTTGCAGATGCCAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246

Db        184 GCCATCGAGTTGGAGCCTACTTTGGCCAAAGCTTATCTCAGAAAGGGCGCTGCTTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306

Db        244 AAGCTAGAAGAGTATGCTACTGCTAAAGCAGCTCTTGAGAAGGGCGCTTCCGTTGCACCA 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366

Db        304 AATGAATCCAAGTTTGAGAAGATGATAGATGAATGCAACCTTCTCATTGCAGAAGAA--- 360

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426

Db        361 -------------GAGAAAGATTTGGTTCAACAGGTGCCACCGACTTTGCCTTCAAGCTC 407

Qy        427 GATTTCACAAATATGGAGAATA-CACCACCAGTGATAGAACCCCCAAGCAAACCAAAATA 485

Db        408 TACAACACCACTAGCAATCGCAGCTGATGCTCCTCCTGCAGCACCTGCCAAACCCATGTT 467

Qy        486 CAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGT 545

Db        468 CAGACACGAGTTCTACCAGAAGCCGGAAGAAGTGGGGGTGGCAATTTTCGCTAAAGGGAT 527

Qy        546 TCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGT 605

Db        528 CCCAAAGCAGAACGTGAATGTAGAGTTTGGTGACCAGATTCTGAGTGTTGTGATTGATGT 587

Qy        606 TCCTGGTGAAGAACCATACCATTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAA 665

Db        588 TGCTGGAGAAGAAGCTTATCATTTCCAGCCAAGATTGTTCGGGAAGATAATACCAGAGAA 647

Qy        666 ATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGT 725

Db        648 GTGCAGATATGAGGTATTGTCGACCAAAGTTGAGATCCGTCTCGCAAAAGCAGAGATAGT 707

Qy        726 GACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCAC 785

Db        708 CACTTGGGCCTCCCTTGAATA---TGGCAAAGGGCAAGCTCTTCTGCCTAAACCCAATGT 764

Qy        786 GCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGA 845

Db        765 TGCATCAGCAGTGTCTCAGAGGCCGGTGTACCCATCCTCTAAGCCGGGAAAAGACTGGGA 824

Qy        846 TAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGC 905

Db        825 CAAGCTGGAAGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTAGATGGAGATGCAGC 884

Qy        906 ATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCAT 965

Db        885 TATGAACAAGTTTTTCAGCGACATATACCAGAGTGCAGATGAGGACATGAGGCGAGCCAT 944

Qy        966 GATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGG 1025

Db        945 GAACAAATCATTTGCAGAGTCCAATGGGACGGTGCTGTCTACAAACTGGAAAGAGGTTGG 1004

Qy       1026 AGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTA 1085

Db       1005 GACAAAGAAAGTGGAGAGCACTCCACCAGATGGCATGGAGCTSAAGAAGTGGGAGTATTG 1064

Qy       1086 A 1086
              |
Db       1065 A 1065


RESULT 13
AJ620882
LOCUS       AJ620882                1068 bp    mRNA    linear   PLN 16-JAN-2004
DEFINITION  Brassica oleracea mRNA for SGT1-like protein (sgt1b-1 gene).
ACCESSION   AJ620882
VERSION     AJ620882.1  GI:40974914
KEYWORDS    SGT1-like protein; sgt1b-1 gene.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Yemm,A.A., Holub,E. and Tor,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1068)
  AUTHORS   Yemm,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2004) Yemm A.A., Sustainable Disease Research,
            Horticultural Research International, Wellesbourne, Warwick, UNITED
            KINGDOM
FEATURES             Location/Qualifiers
     source          1..1068
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /variety="Senna"
                     /db_xref="taxon:3712"
                     /tissue_type="leaf"
     gene            1..1068
                     /gene="sgt1b-1"
     CDS             1..1068
                     /gene="sgt1b-1"
                     /function="role in disease resistance"
                     /codon_start=1
                     /evidence=experimental
                     /product="SGT1-like protein"
                     /protein_id="CAF06580.1"
                     /db_xref="GI:40974915"
```

```
                  /translation="MASELAEKAKEAFLEDDFDVAVDLYSRAIDLDPNCAAFFADRAQ
                  ANIKILNFTEAVADANKAIELEPTLSKAYLRKGTACMKLEEYSTAKAALQKGASVAPN
                  ESKFNKLIDECNLHIAEEEKDLAQQMPPTLPSSSTTPPLATAADSPPAAPAKPMFRHE
                  FYQKPEEVVVTVFAKGIPKQNLNVEFGDQILSVVIDVAGEEAYHFQPRLFGKIIPDKC
                  RYEVLSTKVEIRLAKAEIITWASLEYVKGQALLPKPNVASAVSQRPVYPSSKPAKDWD
                  KLEAEVKKQEKDEKLDGDAAMNKFFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWKE
                  VGTKKVESTPPDGMELKKWEY"
ORIGIN

  Query Match             37.5%;  Score 407.2;  DB 8;  Length 1068;
  Best Local Similarity   63.1%;  Pred. No. 4.8e-63;
  Matches  682;  Conservative    0;  Mismatches  383;  Indels   15;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db          4 GCGAGCGAACTAGCGGAAAAAGCTAAAGAGGCTTTCTTAGAGGACGACTTCGATGTCGCC 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db         64 GTTGACTTATACTCCAGAGCCATTGACTTGGATCCCAATTGCGCCGCCTTCTTCGCCGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db        124 CGTGCTCAGGCCAACATCAAAATCCTTAACTTTACCGAAGCTGTTGCAGATGCCAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db        184 GCCATTGAGTTGGAGCCTACGTTGTCTAAAGCTTACCTCAGAAAAGGCACTGCTTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db        244 AAGCTAGAGGAATACAGTACTGCTAAAGCAGCCCTTCAAAAGGGCGCTTCCGTTGCACCA 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db        304 AATGAATCAAAGTTTAACAAGTTGATAGATGAATGCAATCTTCACATTGCAGAAGAAGAG 363

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db        364 AAAGATTTGGCTCAACAGA------TGCCGCCCACCTTGCCTTCAAGCTCTACCACTCCA 417

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db        418 CCACTAGCAACTGCAGCTGATTCTCCTCCTG------CAGCACCTGCCAAACCCATGTTC 471

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db        472 AGACACGAGTTCTACCAGAAGCCGGAAGAAGTGGTGGTGACAGTTTTCGCTAAAGGGATA 531

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db        532 CCAAAGCAGAACTTGAATGTCGAGTTTGGTGATCAGATTCTGAGTGTTGTGATTGATGTT 591

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db        592 GCTGGAGAGGAAGCTTATCACTTCCAGCCAAGATTGTTTGGGAAGATAATACCAGACAAG 651

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db        652 TGCAGATATGAAGTATTGTCAACCAAAGTTGAGATCCGTCTCGCAAAAGCAGAGATAATC 711

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db        712 ACTTGGGCCTCCCTTGAATA---TGTCAAAGGGCAAGCTCTTCTGCCTAAACCCAATGTC 768

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db        769 GCATCAGCGGTTTCGCAGAGGCCGGTGTACCCATCATCTAAGCCGGCGAAAGACTGGGAC 828

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db        829 AAGCTGGAGGCTGAAGTGAAGAAACAGGAGAAGGATGAGAAGCTAGATGGAGATGCAGCT 888

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db        889 ATGAACAAGTTTTCAGCGATATATACCAGAGTGCTGATGAGGATATGAGGCGAGCCATG 948

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db        949 AACAAATCATTTGCTGAGTCAAATGGGACAGTGCTGTCTACGAACTGGAAAGAGGTTGGG 1008

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db       1009 ACTAAGAAAGTGGAGAGCACTCCACCAGATGGTATGGAGCTCAAGAAGTGGGAGTATTGA 1068


RESULT 14
AF439975
LOCUS       AF439975                1291 bp    mRNA    linear   PLN 06-NOV-2003
DEFINITION  Arabidopsis thaliana SGT1a (SGT1a) mRNA, complete cds.
ACCESSION   AF439975
VERSION     AF439975.1  GI:17017307
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1291)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     The RAR1 interactor SGT1, an essential component of R
            gene-triggered disease resistance
  JOURNAL   Science 295 (5562), 2073-2076 (2002)
  MEDLINE   21893744
   PUBMED   11847307
REFERENCE   2  (bases 1 to 1291)
  AUTHORS   Azevedo,C., Sadanandom,A., Kitagawa,K., Freialdenhoven,A.,
            Shirasu,K. and Schulze-Lefert,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2001) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich NR47UH, UK
FEATURES             Location/Qualifiers
     source          1..1291
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1..1291
                     /gene="SGT1a"
     CDS             77..1129
                     /gene="SGT1a"
                     /codon_start=1
                     /product="SGT1a"
                     /protein_id="AAL33611.1"
                     /db_xref="GI:17017308"
                     /translation="MAKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQ
                     AYIKLESFTEAVADANKAIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPS
                     ESKFKKLIDECNFLITEEEKDLVQPVPSTLPSSVTAPPVSELDVTPTAKYRHEYYQKP
                     EEVVVTVFAKGIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVL
                     STKIEIRLAKADIITWASLEHGKGPAVLPKPNVSSEVSQRPAYPSSKKVKDWDKLEAE
                     VKKQEKDEKLEGDAALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKT
                     IESTPPDGMELKKWEI"
ORIGIN

  Query Match             36.4%;  Score 394.8;  DB 8;  Length 1291;
  Best Local Similarity   62.8%;  Pred. No. 8e-61;
  Matches  678;  Conservative    0;  Mismatches  372;  Indels   30;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db         80 GCGAAGGAGCTTGCTGATAAGGCTAAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT 139

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db        140 GTTGACTTGTACTCCAAAGCCATTGACTTGGATCCTAATTGCGCTGAGTTCTTCGCTGAT 199
```



```
RESULT 15
AY096395
LOCUS       AY096395               1084 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana putative phosphatase (At4g23570) mRNA,
            complete cds.
ACCESSION   AY096395
VERSION     AY096395.1  GI:20465860
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1084)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1084)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
            to this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
            /PGEC) contributed equally to this work as PIs.

            Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .1084
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="U11621"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1. .1084
                     /gene="At4g23570"
     CDS             1. .1053
                     /gene="At4g23570"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative phosphatase"
                     /protein_id="AAM20035.1"
                     /db_xref="GI:20465861"
                     /translation="MAKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQ
                     AYIKLESFTEAVADANKAIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPS
                     ESKFKKLIDECNFLITEEEKDLVQPVPSTLPSSVTAPPVSELDVTPTAKYRHEYYQKP
                     EEVVVTVFAKGIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVL
                     STKIEICLAKADIITWASLEHGKGPAVLPKPNVSSEVSQRPAYPSSKKVKDWDKLEAE
                     VKKQEKDEKLEGDAALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKT
```

```
                 IESTPPDGMELKKWEI"
     3'UTR       1054. .1084
                 /gene="At4g23570"
ORIGIN

  Query Match           36.2%;  Score 393.2;  DB 8;  Length 1084;
  Best Local Similarity  62.7%;  Pred. No. 1.5e-60;
  Matches  677;  Conservative    0;  Mismatches  373;  Indels   30;  Gaps    3;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
              |||  ||| || |      ||||| || ||||| ||||| ||||| ||||||||  | ||
Db          4 GCGAAGGAGCTTGCTGATAAGGCTAAAGAGGCTTTCGTAGACGATGACTTCGATGTTGCT 63

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
                ||  | |||. | | ||||| |||   |  ||       ||| ||  |||  || ||
Db         64 GTTGACTTGTACTCCAAAGCCATTGACTTGGATCCTAATTGCGCTGAGTTCTTCGCTGAT 123

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
              || || |||||  | ||||| ||||  | || |||||||| ||||| ||||| ||||||
Db        124 CGTGCTCAGGCCTATATCAAACTCGAAAGCTTCACTGAGGCCGTGGCAGATGCGAACAAA 183

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
              |||||||| | ||||||     ||   |||||||||     | |||||  |||| || ||
Db        184 GCAATTGAGTTGGATCCTTCATTGACCAAAGCTTACCTAAGAAAAGGAACTGCCTGTATG 243

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
              |||||||||| || |  ||||| ||  | ||||||||    |||| | |||    |  |
Db        244 AAGCTTGAAGAGTATCGGACTGCTAAAACAGCTCTTGAAAAGGGTGCCTCTATCACGCCG 303

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
              | || || | |||        |  |  | |||||  ||    |  ||| | || |||
Db        304 AGTGAATCCAAATTTAAGAAGTTGATAGATGAATGCAATTTTCTAATCACAGAAGAA--- 360

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
                      | | |   |||   | || |  ||         |||    ||||
Db        361 ------------GAGAAAGATTTGGTTCAACCGGTGCCTTCGACTTTGCCTTCAAGTGT 407

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
              ||   ||| |            ||  | | |   | ||   | |    | | | || |||
Db        408 GACAGCACCA-----------CCAGTATCTGAACTTGATGTCACCCCTACAGCCAAGTAC 456

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
              ||||| || |||||| | |    |     ||||| ||  | ||  | |||| || ||  |
Db        457 AGGCACGAGTACTACCAAAAGCCAGAGGAAGTTGTGGTAACTGTTTTTTGCAAAAGGAATA 516

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
              ||      | | |||   ||| || |||||||||||  ||  |||||||    ||||| |||
Db        517 CCCAAGCAGAATGTTAACATCGACTTTGGTGAACAAATTCTGAGTGTTGTGATTGAGGTT 576

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
              ||||| || ||  | ||  || | || ||  |  |||||   ||||| || || || ||
Db        577 CCTGGAGAGGATGCGTATTATCTCCAACCGAGATTGTTTGGAAAGATAATACCAGATAAG 636

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
              ||||||||| |||| || || ||||| |||| ||  | ||||||||||| ||      |
Db        637 TGCAAATATGAAGTATTGTCGACCAAAATTGAGATCTGTCTTGCAAAAGCCGATATAATC 696

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
              ||||||  |  |||| ||  |  |   | | |||  |  | | ||  ||   |  | ||
Db        697 ACATGGGCCTCCCTCGAACACGGCAAAGGGCCAGCGGTTTTGCCAAAGCCTAATGTCTCA 756

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
              |||  |   ||   |  || ||  | || |||||||| |||       || |||||||
Db        757 TCAGAGGTTTCA---CAGAGGCCAGCTTATCCATCTTCTAAGAAAGTTAAGGACTGGGAC 813

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
              || |||||||||||||| || ||| |||| || || || || || || || || || ||
Db        814 AAGCTGGAAGCTGAAGTGAAGAAGCAGGAAAAAGATGAGAAGCTGGAAGGAGACGCAGCT 873

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
              |||||||||||||||| || || ||| || |||| ||||| |||||| || | || |||
Db        874 TTGAACAAATTCTTCCGGGAGATATACCAGAATGCGGATGAGGATATGAGGCGAGCGATG 933

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
              |  || ||||| |||||||| |||| || || || ||||| || ||| |||| |||||
Db        934 AGCAAATCATTTGTGGAATCGAATGGGACAGTGCTGTCAACAAACTGGCAAGAGGTTGGG 993

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
                | ||||   | ||  | |  || || ||||| |||||||||||||| |||||   || |
Db        994 ACTAAGACAATCGAGAGTACTCCTCCGGATGGCATGGAGCTCAAGAAATGGGAGATCTGA 1053

Search completed: April 23, 2005, 08:51:29
Job time : 4822.69 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 00:00:36 ; Search time 3903.36 Seconds
(without alignments)
10590.311 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcggatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1068.4	98.4	1646	3	AY103953	AY103953 Zea mays
2	698.2	64.3	1131	9	CL959360	CL959360 OsIFCC002
3	652.4	60.1	668	2	AW076274	AW076274 614063D11
4	650.2	59.9	752	6	CD443320	CD443320 EL01N0424
5	647	59.6	756	6	CA184150	CA184150 SCRFST314
6	643.4	59.2	659	2	AW090936	AW090936 614068G07
7	643.4	59.2	772	6	CA258233	CA258233 SCCCRT300
8	641	59.0	759	6	CA245820	CA245820 SCEZFL508
9	631	58.1	666	6	CA215140	CA215140 SCSBAD112
10	608.2	56.0	678	6	CA177070	CA177070 SCMCST105
11	607	55.9	665	6	CA153735	CA153735 SCVPRZ203
12	596.4	54.9	622	6	CA229980	CA229980 SCJFFL3C0
13	583.4	53.7	586	2	AW042392	AW042392 614028B03
14	583.4	53.7	609	6	CA229238	CA229238 SCAGFL302
15	572	52.7	702	6	CA262995	CA262995 SCPRLB202
16	566	52.1	718	6	CA106435	CA106435 SCQGHR101
17	565.2	52.0	594	6	CA098274	CA098274 SCMCCL605
18	562	51.7	574	4	BI644145	BI644145 949022C08
19	557.2	51.3	630	6	CA486265	CA486265 WHE4329_D
20	555.6	51.2	579	2	AW066518	AW066518 660015G08
21	552.8	50.9	680	6	CA128103	CA128103 SCAGLR203
22	550	50.6	550	4	BI679413	BI679413 949001E02
23	550	50.6	914	6	CA084222	CA084222 SCEQAM203
24	545	50.2	588	5	BU499669	BU499669 946178B11
25	541	49.8	601	6	CA176126	CA176126 SCJLST102
26	535.8	49.3	553	5	BU098322	BU098322 946134B07
27	529.4	48.7	612	6	CA234143	CA234143 SCCCCL6C0
28	523	48.2	615	6	CA283412	CA283412 SCSBSD105
29	522.6	48.1	730	2	BE362001	BE362001 DG1_83_H0
30	520.8	48.0	725	6	CA094787	CA094787 SCCCCL401
31	520	47.9	578	6	CD996630	CD996630 QBC6a07.x
32	519.6	47.8	708	6	CA240596	CA240596 SCSBFL406
33	514.8	47.4	1089	7	CK208966	CK208966 FGAS02069
34	512.2	47.2	693	6	CA104899	CA104899 SCJFHR1C0
35	511.6	47.1	606	6	CA083869	CA083869 SCEPAM205
36	508.4	46.8	625	6	CA266503	CA266503 SCAGLB204
37	507.4	46.7	518	2	AW057009	AW057009 660007D01
38	507.4	46.7	862	6	CB655564	CB655564 OSJNEc09B
39	507.2	46.7	652	6	CA102017	CA102017 SCAGHR101
c 40	505.2	46.5	798	6	CA176210	CA176210 SCJLST102
41	500.6	46.1	610	6	CA184034	CA184034 SCQSST311
42	499.4	46.0	643	6	CA251380	CA251380 SCQSFL112
43	498.8	45.9	778	6	CA227695	CA227695 SCJLFL301
44	498.4	45.9	607	6	CA120629	CA120629 SCCCLR108
45	487.4	44.9	792	6	CB657723	CB657723 OSJNEc13E

ALIGNMENTS

RESULT 1
AY103953
LOCUS AY103953 1646 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO098564 mRNA sequence.
ACCESSION AY103953
VERSION AY103953.1 GI:21207031
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1646)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1646)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
FEATURES Location/Qualifiers
source 1. .1646
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635643"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
ORIGIN

Query Match 98.4%; Score 1068.4; DB 3; Length 1646;

Best Local Similarity 99.0%; Pred. No. 2.1e-288;
Matches 1075; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 260

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 320

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 380

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 440

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 500

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 560

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 620

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 680

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 740

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 800

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        801 GAAGTTCCTGGTGAAGAACCGTACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 860

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        861 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 920

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780
              |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        921 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTATTCCCCAGAAGATA 980

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840
              |||||||||||||||||||||||||||||||||||||||||||| | ||||| |||||||
Db        981 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGTCAAAAAAGGAC 1040

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900
              |||||||||||||||||||||| ||||||||||||||||||||||||| ||||||||||||
Db       1041 TGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAAGAAAAACTTGAAGGTGAT 1100

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960
              |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||
Db       1101 GCTGCATTGAACAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGG 1160

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1220

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1221 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1280

Qy       1081 TACTAA 1086
              ||||||
Db       1281 TACTAA 1286
```

RESULT 2
CL959360
LOCUS CL959360 1131 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFCC002721 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL959360
VERSION CL959360.1 GI:52373436
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 1131)
AUTHORS Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES Location/Qualifiers
source 1. .1131
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Query Match 64.3%; Score 698.2; DB 9; Length 1131;
Best Local Similarity 79.0%; Pred. No. 1.8e-184;
Matches 889; Conservative 0; Mismatches 188; Indels 48; Gaps 3;

```
Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         13 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGACGACTTCGAGCTC 72

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||||  ||||| ||||| ||| | ||||||||||||||| ||||| |||
Db         73 GCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCCGAGCTCTACGCC 132

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              |||||||||||||| || |||||||| |||||||||||||||||||| || |||||||||
Db        133 GACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTAGCTGATGCTAAC 192

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              || || ||||| ||||| ||    ||||| || |||||    || ||||| ||||||||
Db        193 AAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAAGGCGCTGCATGT 252

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              ||    || || || || |||||||||||| || |||||||| |||||||  || |  |||
Db        253 ATACGACTGGAGGAGTATCAAACTGCAAAAGCGGCTCTTGAATTGGGTTACTCGTTCGCA 312

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              || || || ||||||||| |||| ||| ||||||| |||||||||||||| ||||||||
Db        313 TCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGCATTGCTGAGGAG 372
```

```
Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGG----------CTACTGTGGCTGCT 408
              ||| | |  || || ||||| | ||| |           |  |||| ||| ||
Db        373 CTTAGTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCGCCCTCTGTTGCTTCT 432

Qy        409 ACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCC 468
              | ||||||| ||||| ||| |||| ||||| |||||||||||| || ||||
Db        433 TTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCAATGGTAGA---- 488

Qy        469 CCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACA 528
              |   || |||||||||||||| |||| |||||||||||| |||||||| ||| |||||
Db        489 --AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAAGTTGTATTGACA 546

Qy        529 ATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTG 588
              || ||||| |||||||||||||||||| | ||| ||  |||||||||||||||| |||||
Db        547 ATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGTGAACAAATGTTA 606

Qy        589 AGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCT 648
              ||||| || |||||||| ||||| || || || |||||||||||||| ||||||||||||
Db        607 AGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCTCGTCTGTTTTCT 666

Qy        649 AAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTT 708
              ||||| |||||||||||| ||| ||| |||||  |||||| |||||||||||| | ||
Db        667 AAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTTGAAATAAGACTG 726

Qy        709 GCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTT 768
              || |||||||| ||| | |||||||   | || |||||| | || ||||||| |||||
Db        727 GCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAACCAAAGGCTGTT 786

Qy        769 CCCCAGAAGATAAGCACGC--------------------------CAGCTGAAACAGCC 801
              || || ||||||| | | |                          |||||||| | |||
Db        787 CCACAAAAGATAATCCCTCCAGTTTTACTGATAGCCCTTCTTTTATCAGCTGAATCGGCC 846

Qy        802 CCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAA 861
              |  || || ||||| || || |||||  | || ||||||||||||||||||||||||||||
Db        847 CAGAGGCCATCATATCCTTCCTCAAAATCCAAGAAAGACTGGGATAAACTGGAAGCTGAA 906

Qy        862 GTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTC 921
              || |||||||||||||||| || || ||||| || |||||||||||||||||||| |||
Db        907 GTTAAAAAGGAGGAGAAGGAGGAGAAGCTTGAAGGCGATGCTGCATTGAACAAATTTTTC 966

Qy        922 CGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTG 981
              |||||||||||||  |||||||||||||| |||||  | || |||||||| || || ||
Db        967 CGTGACATCTACAGTGATGCTGATGAAGACATGCGACGAGCAATGATGAAATCTTTTGTT 1026

Qy        982 GAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAA 1041
              ||||| || || |||||||| || |||||||||||||||||  | ||||||||||
Db       1027 GAATCTAACGGTACTGTTCTGTCGACCAATTGGAAAGATGTTGGCTCGAAGAAGGTAGAG 1086

Qy       1042 GGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              || ||||| |||||||| |||||||| ||||| ||||| ||||||
Db       1087 GGAAGCCCACCTGATGGGATGGAGCTTAAGAAATGGGAGTACTAA 1131
```

RESULT 3
AW076274
LOCUS AW076274 668 bp mRNA linear EST 14-OCT-1999
DEFINITION 614063D11.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA, mRNA sequence.
ACCESSION AW076274
VERSION AW076274.1 GI:6031267
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 668)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614063 row: D column: 11.
FEATURES Location/Qualifiers
source 1. .668
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/note="Organ: root; Vector: pBlueScriptII SK+; Site_1: EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"
ORIGIN

Query Match 60.1%; Score 652.4; DB 2; Length 668;
Best Local Similarity 99.8%; Pred. No. 1.1e-171;
Matches 653; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        397 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 ACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCA 73

Qy        457 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAA 133

Qy        517 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGT 193

Qy        577 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCC 253

Qy        637 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTT 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 CGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTC 313

Qy        697 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGA 373

Qy        757 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATAC 433

Qy        817 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAG 493

Qy        877 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 AAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAG 553

Qy        937 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACT 613

Qy        997 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 GTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCC 667
```

RESULT 4
CD443320
LOCUS CD443320 752 bp mRNA linear EST 03-JUN-2003

```
DEFINITION  EL01N0424G10.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION   CD443320
VERSION     CD443320.1  GI:31358963
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F.,
            Larkins,B., Becraft,P. and Messing,J.
  TITLE     Characterization of the maize endosperm transcriptome and its
            comparison to the rice genome
  JOURNAL   Genome Res. 14 (10), 1932-1937 (2004)
COMMENT     Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W22"
                     /db_xref="taxon:4577"
                     /tissue_type="Endosperm of 7-23DAP"
                     /clone_lib="Endosperm_4"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match           59.9%;  Score 650.2;  DB 6;  Length 752;
  Best Local Similarity  99.5%;  Pred. No. 4.9e-171;
  Matches  652;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 157

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 217

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 277

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 337

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 397

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 457

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 517

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 577

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 637

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 697

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTA 655
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        698 GAAGTTCCTGGGTGAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTA 752


RESULT 5
CA184150
LOCUS       CA184150                 756 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCRFST3142H05.g ST3 Saccharum officinarum cDNA clone SCRFST3142H05
            5', mRNA sequence.
ACCESSION   CA184150
VERSION     CA184150.1  GI:35121633
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 756)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 142  row: H  column: 05
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .756
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCRFST3142H05"
                     /lab_host="DH10B"
                     /clone_lib="ST3"
                     /note="Organ: Fourth apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Fourth apical
                     stalk  internodes of adult plants]. cDNA was prepared from
                     polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match           59.6%;  Score 647;  DB 6;  Length 756;
  Best Local Similarity  95.0%;  Pred. No. 3.9e-170;
  Matches  721;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    5;

Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 62

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              | | |||||  ||||| |||||||||||||||||||||||||||||||||||||||| |||
Db         63 GGCGCCGAG--CTACACCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTACGCC 120
```

```
Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        121 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTTGCGGATGCTAAC 180

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              |||||||||||| |||||||||| |||||||||||| ||||||||||||||||||||||
Db        181 AAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGC 240

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        241 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCA 300

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAGGCGATTCAAGGTTTACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 360

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAG 423
              |||||||||||||||||||||| |||||||| |||||||||||||| ||||||||||||
Db        361 TCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAG 420

Qy        424 GAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAA 483
              ||||||  |||||||||||||||||||| |||||| |||||||||||||||||| |||
Db        421 GAGGATGCCACAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAA 480

Qy        484 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGT 543
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        481 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGT 540

Qy        544 GTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAA 603
              |||||||||||||||||||||  ||||||||||||||| ||||||||||||| |
Db        541 GTTCCTGCTGATAGTGTAGTCATTGGATTTGGTGAACAGATGTTAAGTGTATCCATTGGA 600

Qy        604 GTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAG 663
              || ||||||||||||||||||||| |||||||||||| |||||||| ||||||||||
Db        601 GTCCCTGGTGAAGAACCATACCATTNTCAGCCCCGTCTGGTTTCTAAGAATATCCCTGAG 660

Qy        664 AAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAG 723
              ||||| | |||||||||||||||| | |||| |||||||||||||| |||||||||| |
Db        661 AAATG-ANATATCAAGTCTTATCCAACCAGGTCGAAATACGCCTTGC-AAAGCTGAGC-G 717

Qy        724 GTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAG 762
              |||||||||||| ||||||| ||||||||||||||||||
Db        718 GTGACATGGACACCCCTGGA-TATAGTGGAAGACCAAAG 755
```

```
RESULT 6
AW090936
LOCUS       AW090936                659 bp    mRNA    linear   EST 18-OCT-1999
DEFINITION  614068G07.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW090936
VERSION     AW090936.1  GI:6056546
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614068  row: G  column: 07.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN
```

```
  Query Match           59.2%;  Score 643.4;  DB 2;  Length 659;
  Best Local Similarity   99.8%;  Pred. No. 3.8e-169;
  Matches  644;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        240 ATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTA 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTA 74

Qy        300 TGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGA 134

Qy        360 GGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGA 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGA 194

Qy        420 CAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACC 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACC 254

Qy        480 AAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAA 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAA 314

Qy        540 GGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCAT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCAT 374

Qy        600 TGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCC 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCC 434

Qy        660 TGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGA 719
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        435 TGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGA 494

Qy        720 GCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGAT 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 GCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGAT 554

Qy        780 AAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGA 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGA 614

Qy        840 CTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGA 884
              |||||||||||||||||||||||||||||||||||||||||||||
Db        615 CTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGA 659
```

```
RESULT 7
CA258233
LOCUS       CA258233                772 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  SCCCRT3004D05.g RT3 Saccharum officinarum cDNA clone SCCCRT3004D05
            5', mRNA sequence.
ACCESSION   CA258233
VERSION     CA258233.1  GI:35945956
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 772)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
Plate: 004 row: D column: 05
Seq primer: T7 Promoter Primer.
FEATURES Location/Qualifiers
source 1. .772
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCRT3004D05"
/lab_host="DH10B"
/clone_lib="RT3"
/note="Organ: Root apex from adult plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Root apex from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Query Match 59.2%; Score 643.4; DB 6; Length 772;
Best Local Similarity 96.9%; Pred. No. 4e-169;
Matches 656; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 155

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        156 GCCGCCGAGCTCTACACCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTACGCC 215

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        216 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTTGCGGATGCTAAC 275

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              |||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||||
Db        276 AAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGC 335

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        336 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCA 395

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              |||||||||||||||||| |||| |||||||||||||||||||||||||||| ||||||
Db        396 TCAGGCGATTCAAGGTTTACTCGCCTATTGAAGGAATGTGATGAGCGCATCGCAGAGGAA 455

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAG 423
              |||||||||||||||||||||| ||||||||| |||||||||||||| ||||||||||||
Db        456 TCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAG 515

Qy        424 GAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAA 483
              |||||| |||||||||||||||||||| |||||| |||||||||||||||||||| |||
Db        516 GAGGATGCCACAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAA 575

Qy        484 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGT 543
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        576 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGT 635

Qy        544 GTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAA 603
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        636 GTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAA 695

Qy        604 GTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAG 663
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 GTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAG 755

Qy        664 AAATGCAAATATCAAGT 680
              |||||||||||||||||
Db        756 AAATGCAAATATCAAGT 772
```

RESULT 8
CA245820
LOCUS CA245820 759 bp mRNA linear EST 25-SEP-2003
DEFINITION SCEZFL5084D06.g Saccharum officinarum FL5 Saccharum officinarum cDNA clone SCEZFL5084D06 5', mRNA sequence.
ACCESSION CA245820
VERSION CA245820.1 GI:35324833
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 759)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
Plate: 084 row: D column: 06
Seq primer: T7 Promoter Primer.
FEATURES Location/Qualifiers
source 1. .759
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEZFL5084D06"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL5"
/note="Organ: Developed inflorescence (20cm-long) without rachis; Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Developed inflorescence (20cm-long) without rachis]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Query Match 59.0%; Score 641; DB 6; Length 759;
Best Local Similarity 93.6%; Pred. No. 1.9e-168;

```
  Matches  700;  Conservative    0;  Mismatches   43;  Indels    5;  Gaps    3;

Qy        207 GATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 266
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 61

Qy        267 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCG 326
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||||
Db         62 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTTACTCG 121

Qy        327 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 181

Qy        387 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 446
              ||||||||| ||||||||||||| ||||||||||||||||  ||||||||||||||||
Db        182 GGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATGGAGAA 241

Qy        447 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 506
              |||||| |||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db        242 TACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTACAACAG 301

Qy        507 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 566
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 361

Qy        567 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 626
              |||||||||||||||||||| |||||||||||||||||| ||||||||||||||||||||
Db        362 TGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCA 421

Qy        627 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 686
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        422 TTTTCAGCCCCGTCTGTTTTCTAAGAATATCCCTGAGAAATGCAAATATCAAGTCTTATC 481

Qy        687 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 746
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 541

Qy        747 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 806
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        542 TAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 601

Qy        807 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGG---GATAAACTGGAAGCTGAAGT 863
              |||||||||||||||||||||| ||||||     |   |||||   |||||||||
Db        602 ACCTTCATACCCATCTTCAAAGGCAAAAAAAAAGACTGGNNGATAACTGGGAAGCTGAGTC 661

Qy        864 CAAAAAGGAGGAGAA-GGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCC 922
              |||| ||| ||||| ||||||| |||||||||||||||||  ||||    ||||| |||
Db        662 AAAAAGGGAAGAGAAGGGAAGAANAACTTGATGGTGATGCTTGATTGNNACAATTC-TCC 720

Qy        923 GTGACATCTACAAGGATGCTGATGAAGA 950
              |||||||||| | |||||||||| |
Db        721 GTGACATCTAACAAGATGCTGATGAAAA 748


RESULT 9
CA215140
LOCUS       CA215140                 666 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCSBAD1128H12.g AD1 Saccharum officinarum cDNA clone SCSBAD1128H12
            5', mRNA sequence.
ACCESSION   CA215140
VERSION     CA215140.1  GI:35263122
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 128  row: H  column: 12
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCSBAD1128H12"
                     /lab_host="DH10B"
                     /clone_lib="AD1"
                     /note="Organ: seedlings inoculated with Gluconacetobacter
                     diazotroficans; Vector: pSport1; Site_1: SalI; Site_2:
                     NotI; An unidirectional cDNA library generated from
                     [seedlings inoculated with  Gluconacetobacter
                     diazotroficans]. cDNA was prepared from polyA+ mRNA  using
                     SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match           58.1%;  Score 631;  DB 6;  Length 666;
  Best Local Similarity   96.5%;  Pred. No. 1.2e-165;
  Matches  643;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy        339 ATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTAC 398
              |||||||||| |||||||||||||||||||||||||||||||||||| |||||||| |
Db          1 ATGTGATGAGTGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCC 60

Qy        399 TGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGT 458
              |||||||||||| |||||||||||||||  |||||||||||||||||||||| |||||
Db         61 TGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATGGAGAATACACCGCCAGT 120

Qy        459 GATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGT 518
              | ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        121 GGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTACAACAGTGCCACAGAAGT 180

Qy        519 GGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGA 578
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGA 240

Qy        579 ACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCG 638
              ||||||||| |||||||||||||| |||||||||||||||||||||||||||||||| ||
Db        241 ACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCATTTTCAGCCACG 300

Qy        639 TCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGA 698
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        301 TCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGA 360

Qy        699 AATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACC 758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACC 420

Qy        759 AAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCC 818
              |||| |||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGGCTGTTCCCCAAAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCC 480

Qy        819 ATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAA 878
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAA 540
```

```
Qy        879 GGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGA 938
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        541 GGAAGAANAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACANGGA 600

Qy        939 TGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGT 998
              |||||||||||||||||  |||||||||||||||||||||||||| ||| |||||||
Db        601 TGCTGATGAAGATATGCGGNAGGCCATGATGAAGTCATTCGTGGAATCTAATTGCACTGT 660

Qy        999 TCTCTC 1004
              ||||||
Db        661 TCTCTC 666
```

RESULT 10
CA177070

LOCUS	CA177070 678 bp mRNA linear EST 24-SEP-2003
DEFINITION	SCMCST1057G04.g ST1 Saccharum officinarum cDNA clone SCMCST1057G04 5', mRNA sequence.
ACCESSION	CA177070
VERSION	CA177070.1 GI:35107470
KEYWORDS	EST.
SOURCE	Saccharum officinarum
ORGANISM	Saccharum officinarum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE	1 (bases 1 to 678)
AUTHORS	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE	The libraries that made SUCEST
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br Plate: 057 row: G column: 04 Seq primer: T7 Promoter Primer.

FEATURES Location/Qualifiers

```
     source          1. .678
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCMCST1057G04"
                     /lab_host="DH10B"
                     /clone_lib="ST1"
                     /note="Organ: First apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [First apical
                     stalk internodes of adult plants]. cDNA was prepared from
                     polyA+ mRNA using SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match            56.0%;  Score 608.2;  DB 6;  Length 678;
  Best Local Similarity  95.0%;  Pred. No. 3e-159;
  Matches  628;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy        208 ATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACT 267
              | || |   || ||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ACGCGTCCGGCCTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAACT 77

Qy        268 GCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGT 327
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||||
Db         78 GCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTTACTCGT 137

Qy        328 CTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAT 387
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        138 CTATTGAAGGAATGTGATGAGTGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAAG 197

Qy        388 GTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAAT 447
              ||||||||| |||||||||||||| |||||||||||||||||  ||||||||||||||||
Db        198 GTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATGGAGAAT 257

Qy        448 ACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGT 507
              ||||| |||||| |||||||||||||||||||| ||||||||||||||||||||||||||||
Db        258 ACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTACAACAGT 317

Qy        508 GCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATT 567
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATT 377

Qy        568 GATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCAT 627
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db        378 GATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCATACCAT 437

Qy        628 TTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCC 687
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TTTCAGCCACGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCC 497

Qy        688 ACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTAT 747
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTAT 557

Qy        748 AGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGA 807
              |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||||
Db        558 AGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCCCCCCAGCTGAAACAGCCCCAAGA 617

Qy        808 CCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAA 867
              |||||||||||||||| |||||  |||||||||||||||||||||| |   | |||||
Db        618 CCTTCATACCCATCTTTAAAGGGCAAAAAAGACTGGGATAAACTGGAAACCTGAATCAAA 677

Qy        868 A 868
              |
Db        678 A 678
```

RESULT 11
CA153735

LOCUS	CA153735 665 bp mRNA linear EST 24-SEP-2003
DEFINITION	SCVPRZ2035E06.g RZ2 Saccharum officinarum cDNA clone SCVPRZ2035E06 5', mRNA sequence.
ACCESSION	CA153735
VERSION	CA153735.1 GI:35060111
KEYWORDS	EST.
SOURCE	Saccharum officinarum
ORGANISM	Saccharum officinarum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE	1 (bases 1 to 665)
AUTHORS	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE	The libraries that made SUCEST
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br Plate: 035 row: E column: 06

```
          Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .665
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCVPRZ2035E06"
                     /lab_host="DH10B"
                     /clone_lib="RZ2"
                     /note="Organ: Shoot-root transition zone from young plants
                     (small insert library); Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Shoot-root transition zone  from young plants (small
                     insert library)]. cDNA was prepared from polyA+  mRNA
                     using SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             55.9%;  Score 607;  DB 6;  Length 665;
  Best Local Similarity   96.5%;  Pred. No. 6.4e-159;
  Matches  630;  Conservative    0;  Mismatches   22;  Indels    1;  Gaps    1;

Qy        435 AAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGA 494
              |||||||||||||||||| |||||| |||||||||||||||||| ||||||||||||||
Db          1 AAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGA 60

Qy        495 CTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGA 554
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 CTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGA 120

Qy        555 TAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGA 614
              |||||||||||||||||||||||||||||||| |||||||||||||||| ||||||||
Db        121 TAGTGTAGTCATTGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGA 180

Qy        615 AGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATA 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATA 240

Qy        675 TCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGAC 734
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGAC 300

Qy        735 AACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGA 794
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        301 AACCCTGGATTATAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGA 360

Qy        795 AACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGA 854
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        361 AACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGA 420

Qy        855 AGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAA 914
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        421 AGCTGAAGTCAAAAAGGAGGAGAAGGAAGAANAACTTGATGGTGATGCTGCATTGAACAA 480

Qy        915 ATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTC 540

Qy        975 ATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAA 1034
              |||||||||||| |||||||||||||||||||||||||||||||||||||| |||||
Db        541 ATTCGTGGAATCTAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGATCAAAGGT 600

Qy       1035 GGTAGAA-GGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              ||| ||| |||| ||| ||||||||||||||    || ||||||||||||||
Db        601 GGTGGAAGGGGACCCCTCCTGATGGTATGGAAGCTAANAAGTGGGAATACTAA 653


RESULT 12
CA229980
LOCUS       CA229980                622 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCJFFL3C02C11.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCJFFL3C02C11 5', mRNA sequence.
ACCESSION   CA229980
VERSION     CA229980.1  GI:35293036
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: C02  row: C  column: 11
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .622
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCJFFL3C02C11"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             54.9%;  Score 596.4;  DB 6;  Length 622;
  Best Local Similarity   97.4%;  Pred. No. 5.9e-156;
  Matches  606;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        262 CAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTT 321
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db          1 CAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTT 60

Qy        322 GCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTA 120

Qy        382 AAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATG 441
              ||||| |||||||| |||||||||||||| |||||||||||||||  |||||||||
Db        121 AAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATG 180

Qy        442 GAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 501
              |||||||||||| |||||| |||||||||||||||||||||||||||||||||||||||
Db        181 GAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTAC 240

Qy        502 AACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTA 561
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 AACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTA 300
```

```
Qy        562 GTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCA 621
              ||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        301 GTCATTGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCA 360

Qy        622 TACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTC 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTC 420

Qy        682 TTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTG 741
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTG 480

Qy        742 GATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCC 801
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        481 GATTATAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCC 540

Qy        802 CCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAA 861
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        541 CCAAGACCTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGAA 600

Qy        862 GTCAAAAAGGAGGAGAAGGAAG 883
              |||||||||||||||||||||||
Db        601 GTCAAAAAGGAGGAGAAGGAAG 622
```

```
RESULT 13
AW042392
LOCUS       AW042392                586 bp    mRNA    linear   EST 15-SEP-1999
DEFINITION  614028B03.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
            mRNA sequence.
ACCESSION   AW042392
VERSION     AW042392.1  GI:5901292
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 614028  row: B  column: 03.
FEATURES             Location/Qualifiers
     source          1. .586
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="W23"
                     /db_xref="taxon:4577"
                     /tissue_type="root"
                     /dev_stage="3-4 days old"
                     /lab_host="XLOLR"
                     /clone_lib="614 - root cDNA library from Walbot Lab"
                     /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
                     EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
                     Lab (LM)"
ORIGIN

  Query Match             53.7%;  Score 583.4;  DB 2;  Length 586;
  Best Local Similarity   99.7%;  Pred. No. 2.6e-152;
  Matches  584;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        147 GCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTAT 206
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db          1 GCTCGGCAACTACACTGAGGCTGNGGCGGATGCTAACAAAGCAATTGAGCTTGATCCTAT 60

Qy        207 GATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATTAAGCTTGAAGAATACCAAAC 120

Qy        267 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCG 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCG 180

Qy        327 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAA 240

Qy        387 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 300

Qy        447 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 360

Qy        507 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 420

Qy        567 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 480

Qy        627 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 540

Qy        687 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGG 732
              |||||||||| ||||||||||||||||||||||||||||||||||
Db        541 CACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGG 586
```

```
RESULT 14
CA229238
LOCUS       CA229238                609 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCAGFL3023C08.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCAGFL3023C08 5', mRNA sequence.
ACCESSION   CA229238
VERSION     CA229238.1  GI:35291554
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 609)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 023  row: C  column: 08
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .609
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
```

```
                     /clone="SCAGFL3023C08"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing  inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using  SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match           53.7%;  Score 583.4;  DB 6;  Length 609;
  Best Local Similarity   97.4%;  Pred. No. 2.6e-152;
  Matches  593;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        262 CAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTT 321
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 CAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCATCAGGCGATTCAAGGTTT 60

Qy        322 GCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTA 120

Qy        382 AAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATG 441
              ||||| |||||||||| |||||||||||||| |||||||||||||||||  |||||||||
Db        121 AAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAGGAGGATGCCACAAATATG 180

Qy        442 GAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 501
              |||||||||| |||||| |||||||||||||||||| ||||||||||||||||||||||
Db        181 GAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAATACAGGCATGACTACTAC 240

Qy        502 AACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTA 561
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        241 AACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTA 300

Qy        562 GTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCA 621
              |||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db        301 GTCATTGATTTTGGTGAACAGATGTTAAGTGTATCCATTGAAGTCCCTGGTGAAGAACCA 360

Qy        622 TACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTC 681
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTC 420

Qy        682 TTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTG 741
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTG 480

Qy        742 GATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCC 801
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        481 GATTATAGTGGAAGACCAAAGGCTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCC 540

Qy        802 CCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAAGACTGGGATAAACTGGAAGCTGAA 861
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        541 CCAAGACCTTCATACCCATCTTCAAAGGCAAAAAAAGACTGGGATAAACTGGAAGCTGAA 600

Qy        862 GTCAAAAAG 870
              |||||||||
Db        601 GTCAAAAAG 609


RESULT 15
CA262995
LOCUS       CA262995                 702 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  SCPRLB2028A09.g LB2 Saccharum officinarum cDNA clone SCPRLB2028A09
            5', mRNA sequence.
ACCESSION   CA262995
VERSION     CA262995.1  GI:35955513
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 028  row: A  column: 09
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCPRLB2028A09"
                     /lab_host="DH10B"
                     /clone_lib="LB2"
                     /note="Organ: Lateral buds from plants adult plants
                     growing in greenhouse; Vector: pSport1; Site_1: SalI;
                     Site_2: NotI; An unidirectional cDNA library generated
                     from [Lateral buds from plants  adult plants growing in
                     greenhouse]. cDNA was prepared from polyA+ mRNA  using
                     SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand  cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments  sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector.  Details
                     of each source of RNA and library construction can be
                     obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match           52.7%;  Score 572;  DB 6;  Length 702;
  Best Local Similarity   96.4%;  Pred. No. 4.4e-149;
  Matches  584;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 155

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||| ||||||||||||||||||||||||||||||||||||||| |||
Db        156 GCCGCCGAGCTCTACACCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTACGCC 215

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        216 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTTGCGGATGCTAAC 275

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              |||||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||
Db        276 AAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGCTGCATGC 335

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        336 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTATGCA 395

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        396 TCAGGCGATTCAAGGTTTACTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 455

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAG 423
              |||||||||||||||||||||||| ||||||||| |||||||||||||| ||||||||||
```

```
Db        456 TCTAGCCAGGCACCAGTAAAGAAGGTTGAGGCTCCTGTGGCTGCTACTGTTGAGGACAAG 515

Qy        424 GAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAA 483
              ||||||  |||||||||||||||||||| |||||| |||||||||||||||||| |||
Db        516 GAGGATGCCACAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCTAAA 575

Qy        484 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGT 543
              ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        576 TACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTGCTGACAATATTTGCTAAGGGT 635

Qy        544 GTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAA 603
              |||||||||||||||||| |||||||||||||||||||||| | ||||||||||||||
Db        636 GTTCCTGCTGATAGTGTANTCATTGATTTTGGTGAACAGATGTTAAATGTATCCATTGAA 695

Qy        604 GTTCCT 609
              || |||
Db        696 GTNCCT 701
```

Search completed: April 23, 2005, 05:43:24
Job time : 3906.36 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 23, 2005, 03:11:06 ; Search time 685.535 Seconds
(without alignments)
9617.545 Million cell updates/sec

Title: US-10-609-078-9
Perfect score: 1086
Sequence: 1 atggccgcgtcggatctgga.....tcaagaagtgggaatactaa 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1084.4	99.9	1444	17	US-10-425-114-15440	Sequence 15440, A
	2	1084.4	99.9	1450	17	US-10-425-114-25244	Sequence 25244, A
	3	1084.4	99.9	1755	18	US-10-425-115-18040	Sequence 18040, A
	4	998.2	91.9	1537	18	US-10-767-701-13951	Sequence 13951, A
	5	977.6	90.0	1361	17	US-10-425-114-25845	Sequence 25845, A
	6	977.6	90.0	1490	17	US-10-425-114-22598	Sequence 22598, A
	7	977.6	90.0	1519	18	US-10-425-115-18041	Sequence 18041, A
	8	862	79.4	1649	17	US-10-425-114-24871	Sequence 24871, A
	9	862	79.4	1706	18	US-10-425-115-18039	Sequence 18039, A
c	10	736.6	67.8	2471	18	US-10-739-930-4858	Sequence 4858, Ap
	11	698.2	64.3	1569	18	US-10-437-963-37229	Sequence 37229, A
	12	490	45.1	571	16	US-10-074-473-2	Sequence 2, Appli
	13	476.8	43.9	1691	17	US-10-424-599-65699	Sequence 65699, A
	14	474.4	43.7	1598	17	US-10-424-599-81927	Sequence 81927, A
	15	416.6	38.4	483	17	US-10-424-599-118531	Sequence 118531,
	16	332.8	30.6	1674	17	US-10-424-599-98272	Sequence 98272, A
	17	286	26.3	287	9	US-09-294-093B-571	Sequence 571, App
	18	243.4	22.4	1029	17	US-10-424-599-81933	Sequence 81933, A
	19	221.8	20.4	585	18	US-10-021-323-2492	Sequence 2492, Ap
	20	217.8	20.1	534	18	US-10-021-323-14417	Sequence 14417, A
	21	216.4	19.9	272	9	US-09-294-093B-4704	Sequence 4704, Ap
	22	201.6	18.6	851	17	US-10-424-599-75856	Sequence 75856, A
	23	194.4	17.9	528	18	US-10-021-323-14351	Sequence 14351, A
c	24	183.6	16.9	573	18	US-10-021-323-2546	Sequence 2546, Ap
	25	182.8	16.8	275	18	US-10-425-115-60448	Sequence 60448, A
	26	180.6	16.6	656	17	US-10-424-599-98270	Sequence 98270, A
c	27	175.4	16.2	523	18	US-10-021-323-15961	Sequence 15961, A
	28	156.4	14.4	1264	17	US-10-264-049-918	Sequence 918, App
	29	156.4	14.4	1366	16	US-10-287-218-33	Sequence 33, Appl
	30	156.4	14.4	1366	18	US-10-474-291-33	Sequence 33, Appl
	31	156.4	14.4	1555	18	US-10-335-053-99	Sequence 99, Appl
	32	156.4	14.4	1776	18	US-10-357-930-22330	Sequence 22330, A
	33	156.4	14.4	1776	18	US-10-357-930-28174	Sequence 28174, A
	34	156.4	14.4	2053	10	US-09-983-802-45	Sequence 45, Appl
	35	156.4	14.4	2053	10	US-09-984-490-45	Sequence 45, Appl
	36	156.4	14.4	2053	11	US-09-973-278-82	Sequence 82, Appl
	37	148.6	13.7	291	9	US-09-294-093B-3135	Sequence 3135, Ap
c	38	141.4	13.0	409	9	US-09-770-423-621	Sequence 621, App
	39	129.2	11.9	481	10	US-09-770-961-929	Sequence 929, App
	40	125	11.5	388	16	US-10-074-473-3	Sequence 3, Appli
c	41	106.8	9.8	479	17	US-10-424-599-25942	Sequence 25942, A
c	42	106	9.8	432	17	US-10-424-599-49636	Sequence 49636, A
	43	94.4	8.7	317	17	US-10-424-599-41142	Sequence 41142, A
	44	93.2	8.6	672	17	US-10-260-238-1154	Sequence 1154, Ap
	45	93.2	8.6	672	18	US-10-437-963-43909	Sequence 43909, A

ALIGNMENTS

RESULT 1
US-10-425-114-15440
; Sequence 15440, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15440
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-061-D10_FLI
US-10-425-114-15440

Query Match 99.9%; Score 1084.4; DB 17; Length 1444;
Best Local Similarity 99.9%; Pred. No. 5.3e-313;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 142

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 202

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 262

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 322

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 382

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 442

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 502

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 562

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 622

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 682

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 742

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        743 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 802

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 862

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 922

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 982

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 1042

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1043 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1102

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1103 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1162

Qy       1081 TACTAA 1086
              ||||||
Db       1163 TACTAA 1168
```

RESULT 2

```
US-10-425-114-25244
; Sequence 25244, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25244
;   LENGTH: 1450
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3732-024-C2_FLI
US-10-425-114-25244

  Query Match             99.9%;  Score 1084.4;  DB 17;  Length 1450;
  Best Local Similarity   99.9%;  Pred. No. 5.3e-313;
  Matches 1085;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 148

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 208

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 268

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 328

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 388

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 448

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 508

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 568

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 628

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 688

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 748

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
```

```
Db        749 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 808

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780

Db        809 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 868

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840

Db        869 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 928

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900

Db        929 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 988

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960

Db        989 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 1048

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020

Db       1049 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1108

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080

Db       1109 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1168

Qy       1081 TACTAA 1086

Db       1169 TACTAA 1174
```

RESULT 3
US-10-425-115-18040
; Sequence 18040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18040
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116455C.1
US-10-425-115-18040

Query Match 99.9%; Score 1084.4; DB 18; Length 1755;
Best Local Similarity 99.9%; Pred. No. 5.8e-313;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60

Db        252 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 311

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120

Db        312 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 371

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180

Db        372 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 431

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240

Db        432 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 491

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300

Db        492 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 551

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360

Db        552 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 611

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420

Db        612 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 671

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480

Db        672 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 731

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540

Db        732 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 791

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600

Db        792 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 851

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660

Db        852 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 911

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720

Db        912 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 971

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780

Db        972 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 1031

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840

Db       1032 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 1091

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900

Db       1092 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 1151

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960

Db       1152 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 1211

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020

Db       1212 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1271

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080

Db       1272 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1331

Qy       1081 TACTAA 1086

Db       1332 TACTAA 1337
```

RESULT 4
US-10-767-701-13951
; Sequence 13951, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

```
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13951
;   LENGTH: 1537
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS2718_1
US-10-767-701-13951

  Query Match             91.9%;  Score 998.2;  DB 18;  Length 1537;
  Best Local Similarity   95.3%;  Pred. No. 3.3e-287;
  Matches 1043;  Conservative    0;  Mismatches   43;  Indels    9;  Gaps    1;

Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 153

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              |||||| |||||||||||||| |||||||||||||||| |||||||||||||||||| |||||
Db        154 CTGGCCGCCGAGCTCTACACCCAGGCCATCGACGCGGGGCCCGCCACCGCCGAACTCTAC 213

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 273

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              |||||||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||
Db        274 AACAAAGCAATTGAACTTGATCCTACGATGCATAAAGCCTACTACCGGAAAGGTGCTGCA 333

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        334 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCCTAT 393

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              |||||||||||||||||||| |||||||||||||| |||||||| |||||| ||||||||
Db        394 GCATCAGGCGATTCAAGGTTTACTCGTCTATTGAAAGAATGTGAAGAGCGCATTGCTGAG 453

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTG---------GCTGCTACT 411
              |||||||||||||| |||| |||||||||||||| || |||||         |||||||||
Db        454 GAATCTAGCCAGGCTCCAGCAAAGAATGTTGAGCCTCCTGTGGCTGCTGCTGCTGCTACT 513

Qy        412 ATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCA 471
              |||||||||||||||| || |||||||||||||||||||| |||||| ||||||||||||
Db        514 GTTGAGGACAAGGAGGATGTCGCAAATATGGAGAATACACCGCCAGTGGTAGAACCCCCA 573

Qy        472 AGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATA 531
              |||||||| |||||    ||||||||||||||||||||||||||||||||  ||||||||
Db        574 AGCAAACCTAAATAGTCCCATGACTACTACAACAGTGCCACAGAAGTGGTTTTGACAATA 633

Qy        532 TTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGT 591
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db        634 TTTGCTAAGGGTGTTCCCGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTAAGT 693

Qy        592 GTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAG 651
              || |||||||||| |||||||||||||||||||||||||||||||||||||||||| |||||
Db        694 GTGTCCATTGAAGTCCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTGCTAAG 753

Qy        652 ATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCA 711
              |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db        754 ATTATCCCTGAGAAATGCAAATATCAAGTCTTGTCCACCAAGGTCGAAATACGCCTTGCA 813

Qy        712 AAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCC 771
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||
Db        814 AAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGCGGAAGACCAAAGGCTGTTCCC 873

Qy        772 CAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCG 831
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCA 933

Qy        832 AAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTT 891
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 AAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTT 993

Qy        892 GATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGAT 951
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        994 GATGGTGATGCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGAT 1053

Qy        952 ATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAAT 1011
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1054 ATGCGGAGGGCCATGATGAAGTCATTCGTGGAATCTAATGGCACTGTTCTCTCAACCAAT 1113

Qy       1012 TGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAG 1071
              |||||||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||
Db       1114 TGGAAAGATGTTGGATCAAAGAAGGTGGAAGGGAGCCCTCCTGATGGTATGGAGCTCAAG 1173

Qy       1072 AAGTGGGAATACTAA 1086
              |||||||||||||||
Db       1174 AAGTGGGAATACTAA 1188


RESULT 5
US-10-425-114-25845
; Sequence 25845, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25845
;   LENGTH: 1361
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3960-013-B4_FLI
US-10-425-114-25845

  Query Match             90.0%;  Score 977.6;  DB 17;  Length 1361;
  Best Local Similarity   94.1%;  Pred. No. 4.4e-281;
  Matches 1016;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         99 ATGGCCGCGTCGGATCTGGAGAGCAAAGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 158

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              |||||| |||| |||||||  |||||||||||||||||||||||| ||||| |||||||
Db        159 CTGGCCGCCGAACTCTACACGCAGGCCATCGACGCCGGGCCCGCCACTGCCGATCTCTAC 218

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        219 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCAGATGCT 278

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              |||||||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||
Db        279 AACAAAGCAATTGGTCTTGATCCTACAATGCATAAAGCCTACTACCGGAAAGGTGCTGCA 338

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        339 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTAGGTTCTTCCTAT 398

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||| |||||||||||||||| | |||||||||||||||||||||||| |||| ||||||
Db        399 GCACCAGGCGATTCAAGGTTTACCCGTCTATTGAAGGAATGTGATGAGTGCATTGCTGAG 458
```

```
Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              |||||||||||||||||||| ||||||||| |||||| |||| ||||||||| ||||||||
Db        459 GAATCTAGCCAGGCACCAGCAAAGAATGTCGAGGCTCCTGTAGCTGCTACTGTTGAGGAC 518

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||| || |||||||||| |||||||| |||||| ||||||||||||||||||||
Db        519 AAGGAGGATGTCGCAAATATGGATAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCT 578

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||| |||||||||||||||||||||||||||||||| || ||||||| |||||||
Db        579 AAATATAGGCATGACTACTACAACAGTGCCACAGAAGTGGTCCTCACAATATATGCTAAG 638

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        639 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGATCAGATGTTGAGTGTATCCATT 698

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        699 GAAGTACCTGGTGAAGAACCGTACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 758

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
              |||||||||||| |||||||| |||||||||||||| ||||||||||||||||||||||
Db        759 GAGAAATGCAAGTATCAAGTTTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 818

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780
              |||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        819 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTATTCCCCAGAAGATA 878

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||| |||
Db        879 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGTCAAAAAAGGAC 938

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900
              |||||||||||||||||||| |||||||||||||||||||||||||||||||| ||||||
Db        939 TGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAAGAAAAACTTGAAGGTGAT 998

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960
              |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||
Db        999 GCTGCATTGAACAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGG 1058

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020
              ||||||   |||||||||  |||||| ||||| || ||||||||||||||||||||||||
Db       1059 GCCATGGACAAGTCATTCAGGGAATCTAATGGTACCGTTCTCTCAACCAATTGGAAAGAT 1118

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080
              |||||| |||||| ||| |||| |||||| ||||||||||||||||||||||||||||||
Db       1119 GTTGGATCAAAGACGGTGGAAGCGAGCCCTCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1178
```

RESULT 6
US-10-425-114-22598
; Sequence 22598, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22598
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3591-005-E5_FLI
US-10-425-114-22598

Query Match 90.0%; Score 977.6; DB 17; Length 1490;
Best Local Similarity 94.1%; Pred. No. 4.6e-281;
Matches 1016; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        167 ATGGCCGCGTCGGATCTGGAGAGCAAAGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 226

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              |||||| |||| ||||||| ||||||||||||||||||||||||| ||||| |||||
Db        227 CTGGCCGCCGAACTCTACACGCAGGCCATCGACGCCGGGCCCGCCACTGCCGATCTCTAC 286

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        287 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCAGATGCT 346

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              |||||||||||| |||||||||| |||||||||||| ||||||||||||||||||||||
Db        347 AACAAAGCAATTGGTCTTGATCCTACAATGCATAAAGCCTACTACCGGAAAGGTGCTGCA 406

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db        407 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTAGGTTCTTCCTAT 466

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||| |||||||||||||||| | |||||||||||||||||||||| |||| ||||||
Db        467 GCACCAGGCGATTCAAGGTTTACCCGTCTATTGAAGGAATGTGATGAGTGCATTGCTGAG 526

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              |||||||||||||||||||| ||||||||| |||||| |||| ||||||||| ||||||||
Db        527 GAATCTAGCCAGGCACCAGCAAAGAATGTCGAGGCTCCTGTAGCTGCTACTGTTGAGGAC 586

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||| || |||||||||| |||||||| |||||| ||||||||||||||||||||
Db        587 AAGGAGGATGTCGCAAATATGGATAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCT 646

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||| |||||||||||||||||||||||||||||||| || ||||||| |||||||
Db        647 AAATATAGGCATGACTACTACAACAGTGCCACAGAAGTGGTCCTCACAATATATGCTAAG 706

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        707 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGATCAGATGTTGAGTGTATCCATT 766

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        767 GAAGTACCTGGTGAAGAACCGTACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 826

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
              |||||||||||| |||||||| |||||||||||||| ||||||||||||||||||||||
Db        827 GAGAAATGCAAGTATCAAGTTTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 886

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780
              |||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        887 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTATTCCCCAGAAGATA 946

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||| |||
Db        947 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGTCAAAAAAGGAC 1006

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900
              |||||||||||||||||||| |||||||||||||||||||||||||||||||| ||||||
Db       1007 TGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAAGAAAAACTTGAAGGTGAT 1066

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960
              |||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db       1067 GCTGCATTGAACAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGG 1126

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020
              ||||||   |||||||||  |||||| ||||| || ||||||||||||||||||||||||
Db       1127 GCCATGGACAAGTCATTCAGGGAATCTAATGGTACCGTTCTCTCAACCAATTGGAAAGAT 1186
```

```
Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080
              |||||| ||||||| ||| |||| |||||| ||||||||||||||||||||||||||||||
Db       1187 GTTGGATCAAAGACGGTGGAAGCGAGCCCTCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1246
```

RESULT 7
US-10-425-115-18041
; Sequence 18041, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18041
; LENGTH: 1519
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116456C.1
US-10-425-115-18041

Query Match 90.0%; Score 977.6; DB 18; Length 1519;
Best Local Similarity 94.1%; Pred. No. 4.7e-281;
Matches 1016; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        167 ATGGCCGCGTCGGATCTGGAGAGCAAAGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 226

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              |||||| |||| ||||||| |||||||||||||||||||||||| ||||| |||||
Db        227 CTGGCCGCCGAACTCTACACGCAGGCCATCGACGCCGGGCCCGCCACTGCCGATCTCTAC 286

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        287 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCAGATGCT 346

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
              |||||||||||| |||||||||| |||||||||||| ||||||||||||||||||||||
Db        347 AACAAAGCAATTGGTCTTGATCCTACAATGCATAAAGCCTACTACCGGAAAGGTGCTGCA 406

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db        407 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTAGGTTCTTCCTAT 466

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
              ||| |||||||||||||||| | ||||||||||||||||||||||||| |||| ||||||
Db        467 GCACCAGGCGATTCAAGGTTTACCCGTCTATTGAAGGAATGTGATGAGTGCATTGCTGAG 526

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
              |||||||||||||||||| ||||||||| |||||| |||| ||||||||| |||||||||
Db        527 GAATCTAGCCAGGCACCAGCAAAGAATGTCGAGGCTCCTGTAGCTGCTACTGTTGAGGAC 586

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
              ||||||||| || |||||||||| |||||||| |||||| ||||||||||||||||||||
Db        587 AAGGAGGATGTCGCAAATATGGATAATACACCGCCAGTGGTAGAACCCCCAAGCAAACCT 646

Qy        481 AAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAG 540
              ||||| |||||||||||||||||||||||||||||||||| || ||||||| |||||||
Db        647 AAATATAGGCATGACTACTACAACAGTGCCACAGAAGTGGTCCTCACAATATATGCTAAG 706

Qy        541 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATT 600
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        707 GGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGATCAGATGTTGAGTGTATCCATT 766

Qy        601 GAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 660
              ||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        767 GAAGTACCTGGTGAAGAACCGTACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCT 826

Qy        661 GAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAG 720
              |||||||||| |||||||| |||||||||||||| ||||||||||||||||||||||||
Db        827 GAGAAATGCAAGTATCAAGTTTTATCCACCAAGGTCGAAATACGCCTTGCAAAAGCTGAG 886

Qy        721 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATA 780
              |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        887 CAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGGCTATTCCCCAGAAGATA 946

Qy        781 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGAC 840
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||| |||
Db        947 AGCACGCCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGTCAAAAAAGGAC 1006

Qy        841 TGGGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGAT 900
              |||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||
Db       1007 TGGGATAAACTGGAAGCTGAAGTAAAAAAGGAGGAGAAGGAAGAAAAACTTGAAGGTGAT 1066

Qy        901 GCTGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGG 960
              |||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||
Db       1067 GCTGCATTGAACAAGTTCTTCCGTGATATCTACAAGGATGCTGATGAAGATATGCGGAGG 1126

Qy        961 GCCATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGAT 1020
              ||||||  |||||||||  |||||| ||||| || ||||||||||||||||||||||||
Db       1127 GCCATGGACAAGTCATTCAGGGAATCTAATGGTACCGTTCTCTCAACCAATTGGAAAGAT 1186

Qy       1021 GTTGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1080
              |||||| ||||||| ||| |||| |||||| ||||||||||||||||||||||||||||||
Db       1187 GTTGGATCAAAGACGGTGGAAGCGAGCCCTCCTGATGGTATGGAGCTCAAGAAGTGGGAA 1246
```

RESULT 8
US-10-425-114-24871
; Sequence 24871, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24871
; LENGTH: 1649
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-210-C7_FLI
US-10-425-114-24871

Query Match 79.4%; Score 862; DB 17; Length 1649;
Best Local Similarity 83.5%; Pred. No. 1.8e-246;
Matches 1086; Conservative 0; Mismatches 0; Indels 214; Gaps 1;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 150

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 210

Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 270
```

```
Qy       181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
Db       271 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 330

Qy       241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
Db       331 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 390

Qy       301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCT--- 357
Db       391 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGGT 450

Qy       358 ------------------------------------------------------------ 357
Db       451 GAGAAAAATTAACTTCAGCAGTGTGTGTGTATCTGGTGGTGGTGTTGGGGTCTTGGGTAT 510

Qy       358 ------------------------------------------------------------ 357
Db       511 GCAATGTGGTGCATGCTATTTTCTGACGATATGAGTCCACTTAGCTGCCTTTCCTGAGAA 570

Qy       358 ------------------------------GAGGAATCTAGCCAGGCACCAGTAAAGAA 386
Db       631 TACCTGTCATGTTATTTTGTTTATGTTTCCAGAGGAATCTAGCCAGGCACCAGTAAAGAA 690

Qy       387 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 446
Db       691 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 750

Qy       447 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 506
Db       751 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 810

Qy       507 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 566
Db       811 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 870

Qy       567 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 626
Db       871 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 930

Qy       627 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 686
Db       931 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 990

Qy       687 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 746
Db       991 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 1050

Qy       747 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 806
Db      1051 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 1110

Qy       807 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAA 866
Db      1111 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAA 1170

Qy       867 AAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGA 926
Db      1171 AAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGA 1230

Qy       927 CATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATC 986
Db      1231 CATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATC 1290

Qy       987 AAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAG 1046
Db      1291 AAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAG 1350

Qy      1047 CCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db      1351 CCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1390
```

RESULT 9
US-10-425-115-18039
; Sequence 18039, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 18039
; LENGTH: 1706
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116454C.1
US-10-425-115-18039

Query Match 79.4%; Score 862; DB 18; Length 1706;
Best Local Similarity 83.5%; Pred. No. 1.9e-246;
Matches 1086; Conservative 0; Mismatches 0; Indels 214; Gaps 1;

```
Qy         1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
Db        91 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 150

Qy        61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
Db       151 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 210

Qy       121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
Db       211 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 270

Qy       181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
Db       271 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 330

Qy       241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
Db       331 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 390

Qy       301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCT--- 357
Db       391 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGGT 450

Qy       358 ------------------------------------------------------------ 357
Db       451 GAGAAAAATTAACTTCAGCAGTGTGTGTGTATCTGGTGGTGGTGTTGGGGTCTTGGGTAT 510

Qy       358 ------------------------------------------------------------ 357
Db       511 GCAATGTGGTGCATGCTATTTTCTGACGATATGAGTCCACTTAGCTGCCTTTCCTGAGAA 570

Qy       358 ------------------------------------------------------------ 357
Db       571 ATGTTTGTAGCAGCCTTGTAGACTGGGGGATGATTATGTTTCAGCTTAATCATCTCATAC 630

Qy       358 ------------------------------GAGGAATCTAGCCAGGCACCAGTAAAGAA 386
Db       631 TACCTGTCATGTTATTTTGTTTATGTTTCCAGAGGAATCTAGCCAGGCACCAGTAAAGAA 690

Qy       387 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 446
```

```
Db        691 TGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAA 750

Qy        447 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 506

Db        751 TACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAG 810

Qy        507 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 566

Db        811 TGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCAT 870

Qy        567 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 626

Db        871 TGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCA 930

Qy        627 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 686

Db        931 TTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATC 990

Qy        687 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 746

Db        991 CACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTA 1050

Qy        747 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 806

Db       1051 TAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCCCCAAG 1110

Qy        807 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAA 866

Db       1111 ACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAAGTCAA 1170

Qy        867 AAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGA 926

Db       1171 AAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTCCGTGA 1230

Qy        927 CATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATC 986

Db       1231 CATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTGGAATC 1290

Qy        987 AAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAG 1046

Db       1291 AAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAAGGGAG 1350

Qy       1047 CCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086

Db       1351 CCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1390
```

```
RESULT 10
US-10-739-930-4858/c
; Sequence 4858, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 4858
;   LENGTH: 2471
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(2471)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER396_1
US-10-739-930-4858

  Query Match             67.8%;  Score 736.6;  DB 18;  Length 2471;
  Best Local Similarity   80.5%;  Pred. No. 7.3e-209;
```

```
  Matches  906;  Conservative    0;  Mismatches  174;  Indels   45;  Gaps    2;

Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63

Db       2279 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 2220

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123

Db       2219 GCCGCCGAGCTCTACACCCAGGCCATTGAGGCGGGGCCGGCCACCGCGGAACTCTACGCC 2160

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183

Db       2159 GACCGAGCCCAGGCGCACATCAAGCTGGGCAGTTACACTGAGGCTGTAGCTGATGCCAAC 2100

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243

Db       2099 AAAGCAATTGAACTTGATCCTTCGATGCATAAAGCATACCTTCGGAAGGGCTCTGCTTGC 2040

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303

Db       2039 ATCAAGCTGGAGGAATACCAAACTGCAAAGGCTGCTCTTGAAGTGGGTTCTTCTTATGCA 1980

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363

Db       1979 TCTGGTGACTCGAGGTTTACTCGTCTGATGAAGGAGTGTGATGATCGTATTGCTGAGGAG 1920

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCT------------------ 405

Db       1919 GCTAGCCAGGTGCCAGTAAAGAATGCCGCTGCGGCTGTTGCTTCAGCTACATCTTCGGGG 1860

Qy        406 ------------------------GCTACTATTGAGGACAAGGAGGATTTCACAAATATG 441

Db       1859 GCATCTTCCGGGGCTACAACTGTGGCTACTGAAGCTGAGGACCAGGATGGTGCAAATATG 1800

Qy        442 GAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATACAGGCATGACTACTAC 501

Db       1799 GAGAATGCACAGCCAACGATAGAAGTGCCAAGCAAGCCCAAATACAGGCATGACTACTAC 1740

Qy        502 AACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTA 561

Db       1739 AATACTCCTACAGAAGTGGTACTGACTATATTTGCTAAGGGTGTTCCAGCTGACAGCGTG 1680

Qy        562 GTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCA 621

Db       1679 GTTGTTGACTTTGGTGAACAGATGCTGAGTGTCTCAATTGAACTTCCCGGTGAGGAACCA 1620

Qy        622 TACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTC 681

Db       1619 TACCATTTTCAGCCTCGTCTGTTTTCAAAGATCGTCCCAGATAAGTGCAAGTATACTGTG 1560

Qy        682 TTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTG 741

Db       1559 TTGTCTACAAAGGTTGAAATGCGCCTTGCAAAAGCTGAGCCAGTAACTTGGACATCATTG 1500

Qy        742 GATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACGCCAGCTGAAACAGCC 801

Db       1499 GATTATACTGGTAAACCAAAG---GCTCCTCAGAAGATAAATGTACCAGCTGAATCAGCC 1443

Qy        802 CCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAA 861

Db       1442 CAGAGGCCATCTTATCCATCATCAAAATCCAAAAAGGACTGGGATAAGCTTGAAGCTGAA 1383

Qy        862 GTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTC 921

Db       1382 GTGAAAAAACAGGAGAAGGATGAAAAACTTGATGGCGATGCTGCATTGAACAAATTTTTC 1323

Qy        922 CGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTG 981

Db       1322 CGTGAAATTTACAGTGATGCTGATGAAGATATGCGTAGAGCAATGATGAAGTCTTTTGTG 1263

Qy        982 GAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAA 1041

Db       1262 GAGTCTAATGGAACCGTTCTCTCAACCAACTGGAAGGATGTCGGGAAAAAGACGGTTGAA 1203
```

```
Qy       1042 GGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              || ||||| |||||||| |||||||||||||||||||| || |||
Db       1202 GGAAGCCCTCCTGATGGAATGGAGCTCAAGAAGTGGGAGTATTAA 1158
```

RESULT 11
US-10-437-963-37229
; Sequence 37229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37229
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40977C.1
US-10-437-963-37229

Query Match 64.3%; Score 698.2; DB 18; Length 1569;
Best Local Similarity 79.0%; Pred. No. 1.7e-197;
Matches 889; Conservative 0; Mismatches 188; Indels 48; Gaps 3;

```
Qy          4 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTG 63
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         94 GCCGCGTCGGATCTGGAGAGCAAGGCCAAGGCGGCCTTCGTCGACGACGACTTCGAGCTC 153

Qy         64 GCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCC 123
              ||| |||||||||||| ||||| ||||| ||| | |||||||||||||| ||||| |||
Db        154 GCCGCCGAGCTCTACACGCAGGCAATCGAGGCCAGCCCCGCCACCGCCGAGCTCTACGCC 213

Qy        124 GACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAAC 183
              |||||||||||||| || |||||||| |||||||||||||||||||| || |||||||||
Db        214 GACCGCGCCCAGGCCCATATCAAGCTAGGCAACTACACTGAGGCTGTAGCTGATGCTAAC 273

Qy        184 AAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGC 243
              || || ||||| ||||| ||    ||||| || |||||    || ||||| |||||||||
Db        274 AAGGCCATTGAACTTGACCCATCAATGCACAAGGCTTATCTTCGTAAAGGCGCTGCATGT 333

Qy        244 ATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCA 303
              ||    || || || |||||||||| || |||||||| |||||||  || |  |||
Db        334 ATACGACTGGAGGAGTATCAAACTGCAAAAGCAGCTCTTGAATTGGGTTACTCGTTCGCA 393

Qy        304 TCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAA 363
              || || || ||||||||| |||| ||| ||||||| |||||||||||||| ||||||||
Db        394 TCTGGTGACTCAAGGTTTACTCGCCTAATGAAGGAGTGTGATGAGCGCATTGCTGAGGAG 453

Qy        364 TCTAGCCAGGCACCAGTAAAGAATGTTGAGG----------------CTACTGTGGCTGCT 408
                |||  | |  || || ||||| | ||| |                |  |||| ||| ||
Db        454 CTTAGTGAAGTCCCTGTTAAGAAGGCTGAAGATGGAGCAGCTGCCCCCTCTGTTGCTTCT 513

Qy        409 ACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCC 468
              | ||||||| ||||| |||    |||| ||||| |||||||||||| || ||||
Db        514 TTTGTTGAGGAAAAGGATGATGCTGCAAACATGGATAATACACCACCAATGGTAGA---- 569

Qy        469 CCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACA 528
                |   || |||||||||||||| |||| |||||||||||| |||||||| ||| |||||
Db        570 --AGTGAAGCCAAAATACAGGCACGACTTCTACAACAGTGCTACAGAAGTTGTATTGACA 627

```
Qy   529 ATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTG 588
              || ||||| |||||||||||||||||| | ||| || |||||||||||||||| |||||
Db   628 ATTTTTGCAAAGGGTGTTCCTGCTGAGAATGTTGTTGTTGATTTTGGTGAACAAATGTTA 687

Qy   589 AGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCT 648
              ||||| || |||||||| ||||| || || || |||||||||||||| ||||||||||||
Db   688 AGTGTGTCGATTGAAGTCCCTGGAGAGGAGCCGTACCATTTTCAGCCTCGTCTGTTTTCT 747

Qy   649 AAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTT 708
              ||||| |||||||||||| ||| ||| ||||| ||||||| |||||||||||| | ||
Db   748 AAGATCATCCCTGAGAAAAGCAGATACCAAGTGCTATCCACGAAGGTTGAAATAAGACTG 807

Qy   709 GCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTT 768
              || |||||||| ||| | |||||||| | || |||||| | || ||||||| |||||
Db   808 GCTAAAGCTGAACAGATTACATGGACCTCACTTGATTATGATAAAAAACCAAAGGCTGTT 867

Qy   769 CCCCAGAAGATAAGCACGC--------------------------CAGCTGAAACAGCC 801
              || || ||||||| | | | |||||||| | |||
Db   868 CCACAAAAGATAATCCCTCCAGTTTTACTGATAGCCCTTCTTTTATCAGCTGAATCGGCC 927

Qy   802 CCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGATAAACTGGAAGCTGAA 861
              | || || ||||| || || ||||| | || ||||||||||||||||||||||||||||
Db   928 CAGAGGCCATCATATCCTTCCTCAAAATCCAAGAAAGACTGGGATAAACTGGAAGCTGAA 987

Qy   862 GTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCATTGAACAAATTCTTC 921
              || |||||||||||||||| || || ||||| || |||||||||||||||||||| |||
Db   988 GTTAAAAAGGAGGAGAAGGAGGAGAAGCTTGAAGGCGATGCTGCATTGAACAAATTTTTC 1047

Qy   922 CGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATGATGAAGTCATTCGTG 981
              ||||||||||||| |||||||||||||| ||||| | || |||||||| || || ||
Db   1048 CGTGACATCTACAGTGATGCTGATGAAGACATGCGACGAGCAATGATGAAATCTTTTGTT 1107

Qy   982 GAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGAGCAAAGAAGGTAGAA 1041
              ||||| || || |||||||| || ||||||||||||||||||| | ||||||||||
Db   1108 GAATCTAACGGTACTGTTCTGTCGACCAATTGGAAAGATGTTGGCTCGAAGAAGGTAGAG 1167

Qy   1042 GGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
              || ||||| |||||||| |||||||| ||||| ||||| ||||||
Db   1168 GGAAGCCCACCTGATGGGATGGAGCTTAAGAAATGGGAGTACTAA 1212

RESULT 12
US-10-074-473-2
; Sequence 2, Application US/10074473
; Publication No. US20030167504A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Edmund H. III
; TITLE OF INVENTION: Maize Rar1 Interactor Polynucleotides
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: 35718/241314
; CURRENT APPLICATION NUMBER: US/10/074,473
; CURRENT FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/268,157
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Zea mays
US-10-074-473-2

Query Match 45.1%; Score 490; DB 16; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.7e-135;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ATGGCCGCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAG 141

Qy         61 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CTGGCCACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTAT 201
```

```
Qy        121 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 180
Db        202 GCCGACCGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCT 261

Qy        181 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 240
Db        262 AACAAAGCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCA 321

Qy        241 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 300
Db        322 TGCATTAAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTAT 381

Qy        301 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 360
Db        382 GCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAG 441

Qy        361 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 420
Db        442 GAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGAC 501

Qy        421 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 480
Db        502 AAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCA 561

Qy        481 AAATACAGGC 490
Db        562 AAATACAGGC 571
```

RESULT 13
US-10-424-599-65699
; Sequence 65699, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65699
; LENGTH: 1691
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_30341C.1
US-10-424-599-65699

```
  Query Match             43.9%;  Score 476.8;  DB 17;  Length 1691;
  Best Local Similarity   65.7%;  Pred. No. 2.8e-131;
  Matches  710;  Conservative    0;  Mismatches  367;  Indels    3;  Gaps    1;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db        178 GCTTCCGATCTCGAAGCAAAAGCCAAAGAGGCCTTCGTCGAAGACCACTTCGAACTCGCC 237

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db        238 GTGGACCTTCTCTCTCAGGCCATTCATCTCGAACCCAACAAAGCCGAATTCTACGCCGAC 297

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db        298 CGCGCCCAAGCCAACATCAAACTCAACAACTTCACTGAGGCTGTGGCTGATGCTAACAAA 357

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db        358 GCAATTGAATTGAATCCTTCTCTACCAAAAGCATATTTGCGGAAAGGTACTGCTTGCATG 417

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db        418 AAGCTTGAGGAATATGAAACTGCAAAGGCTGCTCTGGAGGTGGGTGCATCATTGTCTCCA 477

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db        478 GACAACTCAAGGTTTGCTACTTTGATCAAAGAATGTGATAAGCTCATTGCAGAAGAATCT 537

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db        538 TATACCCCACCTATTATAGAGGAAAAGGCCACAACAGAAGATGTTACT---CCAAAAGAT 594

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db        595 GTTCAGCAACAAGATGATCTTCTAGAGAAGCCAACAGTTACAGTGACTAAACCTAAATAT 654

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db        655 AGGCATGAATTCTACCAGAAACATGATCAGGTGGTTGTGACCATATTTGCAAAGAAAATC 714

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db        715 CCAAAGGAAAGCATTACTGTTGACTTCGGTGAACAAATACTAAGTGTTAGTATTAATGTC 774

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db        775 ACAGGAGAAGACACATATGTTTTTCAACCTCGCTTATTTGGAAAGATCATACCTTCCAAT 834

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db        835 TGTCAATATGAAGTTTTGTCCACCAAAATTGAAATTCGTCTTGCAAAAGCAGAACCTATT 894

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db        895 CACTGGACATCTCTAGAATTCACCAGAGATAGCGTTGTTCAACAGAGGGTTAGTGCTTTG 954

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAAAGACTGGGAT 846
Db        955 TCAGTTACTGGAGGTGAAAAACCTACTTATCCATCCTCAAAACGCTGCGAAAATACGACA 1014

Qy        847 AAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGCTGCA 906
Db       1015 GAAGGGGAAGCTCAAGTCAAGAAAGAGGGGAAAGATGAAAAGCTTGATGGCGATGCTGCA 1074

Qy        907 TTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGCCATG 966
Db       1075 TTGAACAAATTTTTCCGCGAAATATATCAAGATGCAGATGAGGATACAAGAAGAGCAATG 1134

Qy        967 ATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGTTGGA 1026
Db       1135 AAAAAGTCATTTGTTGAGTCTAATGGGACAGTTCTGTCTACAAACTGGAAAGAAGTGGGA 1194

Qy       1027 GCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATACTAA 1086
Db       1195 ACAAAGAAGGTAGAGGGAAGTCCTCCTGATGGCATGGAGTTGAAGAAATGGGAATATTAA 1254
```

RESULT 14
US-10-424-599-81927
; Sequence 81927, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 81927

```
;   LENGTH: 1598
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_44998C.1
US-10-424-599-81927

  Query Match             43.7%;  Score 474.4;  DB 17;  Length 1598;
  Best Local Similarity   66.2%;  Pred. No. 1.4e-130;
  Matches  718;  Conservative    0;  Mismatches  356;  Indels   10;  Gaps    2;

Qy          7 GCGTCGGATCTGGAGAGCAAGGCCAAGGAGGCCTTCGTCGACGACGACTTCGAGCTGGCC 66
Db        145 GCTTCCGATCTTGAGCTTAAGGCCAAAGAGGCTTTCGAAGATGATAACTATGATCTCGCC 204

Qy         67 ACCGAGCTCTACAGCCAGGCCATCGACGCCGGGCCCGCCACCGCCGACCTCTATGCCGAC 126
Db        205 TACGACCTCTTAACTCAGGCCATTGGTCTCAGCCCCAACAACGCAGACCTATATGCTGAC 264

Qy        127 CGCGCCCAGGCGCACATCAAGCTCGGCAACTACACTGAGGCTGTGGCGGATGCTAACAAA 186
Db        265 CGTGCGCAAGTCAACATCAAAGTCAACAACCTCACCGAGGCTGTTTCTGATGCAAACAAG 324

Qy        187 GCAATTGAGCTTGATCCTATGATGCATAAAGCTTACTACCGGAAAGGTGCTGCATGCATT 246
Db        325 GCGATTGAGTTGAATCCTTCTCACTCAAAAGCATATTTGCGAAAAGGTACCGCATGCATC 384

Qy        247 AAGCTTGAAGAATACCAAACTGCAAAGGCTGCTCTTGAGTTGGGTTCTTCTTATGCATCA 306
Db        385 AAGCTTGAGGAATATCAGACTGCTAAGGCAGCTCTAGAGATGGGTGCTTCATTGGCTCCT 444

Qy        307 GGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGTGATGAGCGCATCGCTGAGGAATCT 366
Db        445 GGAGATTCTAAATTTACTGATTTGATCAAAGACTGCGATGAACTGATTGCAGAAGAATCT 504

Qy        367 AGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTGGCTGCTACTATTGAGGACAAGGAG 426
Db        505 GGTGTCATACCCATACAAGAAGAGAGCACAACACAGGGTGCTGCTACAAAAGCTGTTGAG 564

Qy        427 GATTTCACAAATATGGAGAATACACCACCAGTGATAGAACCCCCAAGCAAACCAAAATAC 486
Db        565 G------CAGAGAATGATCTTCCAGAGCCACCTACAGTAACCGTGGTTAAACCTAAATAC 618

Qy        487 AGGCATGACTACTACAACAGTGCCACAGAAGTGGTACTGACAATATTTGCTAAGGGTGTT 546
Db        619 AGGCATGAATTCTACCAGAAACCTGATGAAATGGTTATTACCATATTTGCAAAGGGCATT 678

Qy        547 CCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAGATGTTGAGTGTATCCATTGAAGTT 606
Db        679 CCACGAGACAGCATTACTGTTGACTTTGGTGAACAAATATTAAGTGTTACAATTAATATC 738

Qy        607 CCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTGTTTTCTAAGATTATCCCTGAGAAA 666
Db        739 CCTTGCAAAGATGCTTATGTCTTCCAACCTCGCTTATTTGGAAAGATCATACCCTCCAAA 798

Qy        667 TGCAAATATCAAGTCTTATCCACCAAGGTTGAAATACGCCTTGCAAAAGCTGAGCAGGTG 726
Db        799 TGCCGGTATGAAGTTTTGTCCACCAAAATTGAAATTTGCCTTGCAAAAGCAGATCATATC 858

Qy        727 ACATGGACAACCCTGGATTATAGTGGAAGACCAAAGACTGTTCCCCAGAAGATAAGCACG 786
Db        859 CAATGGACATCTCTAGAATTCAACAAGGGTAGCACAGTTGCACAGAGGTTTAGTGTTTTG 918

Qy        787 CCAGCTGAAACAGCCCCAAGACCTTCATACCCATCTTCAAAGGCGAAAAA----AGACTG 842
Db        919 CCAGTTGCAAGAGGTGAAAAACCTACTTATCCATCCTCAAAACCGAAAAAAATACAGATTG 978

Qy        843 GGATAAACTGGAAGCTGAAGTCAAAAAGGAGGAGAAGGAAGAAAAACTTGATGGTGATGC 902
Db        979 GGATAAGCTTGAAGCTCAAGTTAAGAAAGAGGAGAAAGAAGAAAAACTTGATGGTGATGC 1038

Qy        903 TGCATTGAACAAATTCTTCCGTGACATCTACAAGGATGCTGATGAAGATATGCGGAGGGC 962
Db       1039 TGCGTTGAACAAATTTTTCCGTGATATATATCAAGATGCAGATGAGGACACAAGAAGAGC 1098

Qy        963 CATGATGAAGTCATTCGTGGAATCAAATGGCACTGTTCTCTCAACCAATTGGAAAGATGT 1022
Db       1099 AATGAGCAAATCATTTGTGGAGTCTAATGGAACAGTACTGTCTACAAACTGGAAAGAAGT 1158

Qy       1023 TGGAGCAAAGAAGGTAGAAGGGAGCCCCCCTGATGGTATGGAGCTCAAGAAGTGGGAATA 1082
Db       1159 GGGATCAATGAAGGTACAGGAAAGTCCTCCTGATGGCATGGAGTTAAAGAAATGGTAATA 1218

Qy       1083 CTAA 1086
Db       1219 TTGA 1222

RESULT 15
US-10-424-599-118531
; Sequence 118531, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 118531
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_78041C.1
US-10-424-599-118531

  Query Match             38.4%;  Score 416.6;  DB 17;  Length 483;
  Best Local Similarity   91.9%;  Pred. No. 1.4e-113;
  Matches  440;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        283 GAGTTGGGTTCTTCTTATGCATCAGGCGATTCAAGGTTTGCTCGTCTATTGAAGGAATGT 342
Db          5 GCGTCCGGTTCTTCCTATGCACCAGGCGATTCAAGGTTTACCCGTCTATTGAAGGAATGT 64

Qy        343 GATGAGCGCATCGCTGAGGAATCTAGCCAGGCACCAGTAAAGAATGTTGAGGCTACTGTG 402
Db         65 GATGAGTGCATTGCTGAGGAATCTAGCCAGGCACCAGCAAAGAATGTCGAGGCTCCTGTA 124

Qy        403 GCTGCTACTATTGAGGACAAGGAGGATTTCACAAATATGGAGAATACACCACCAGTGATA 462
Db        125 GCTGCTACTGTTGAGGACAAGGAGGATGTCGCAAATATGGATAATACACCGCCAGTGGTA 184

Qy        463 GAACCCCCAAGCAAACCAAAATACAGGCATGACTACTACAACAGTGCCACAGAAGTGGTA 522
Db        185 GAACCCCCAAGCAAACCTAAATATACGCATGACTACTACAACAGTGCCACAGAAGTGGTC 244

Qy        523 CTGACAATATTTGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTTTGGTGAACAG 582
Db        245 CTCACAATATATGCTAAGGGTGTTCCTGCTGATAGTGTAGTCATTGATTGTGGTGATCAG 304

Qy        583 ATGTTGAGTGTATCCATTGAAGTTCCTGGTGAAGAACCATACCATTTTCAGCCCCGTCTG 642
Db        305 ATGTTGAGTGTATCCATTGAAGTACCTGGTGAAGAACCGTACCATTTTTAGCCCCGTCTG 364

Qy        643 TTTTCTAAGATTATCCCTGAGAAATGCAAATATCAAGTCTTATCCACCAAGGTTGAAATA 702
Db        365 TTATCTAAGATTATCCCTGAGAAATGCAAGTATCACGTTCTATCCACCAAGGTCGAAATA 424

Qy        703 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACCCTGGATTATAGTGGAAGACCAAA 761
Db        425 CGCCTTGCAAAAGCTGAGCAGGTGACATGGACAACTCTTGATTATAGCGGAAGACCAAA 483
```

Search completed: April 23, 2005, 09:26:00
Job time : 688.535 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:26:41 ; Search time 43 Seconds
(without alignments)
626.705 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	464	24.9	196	3	US-09-227-357-393	Sequence 393, App
2	464	24.9	228	3	US-09-227-357-401	Sequence 401, App
3	323	17.4	181	4	US-09-270-767-44320	Sequence 44320, A
4	244	13.1	388	4	US-09-248-796A-18350	Sequence 18350, A
5	206	11.1	65	4	US-09-248-796A-16433	Sequence 16433, A
6	197.5	10.6	1427	4	US-09-551-974A-97	Sequence 97, Appl
7	197.5	10.6	1427	4	US-09-565-501A-97	Sequence 97, Appl
8	197.5	10.6	1427	4	US-09-639-206A-97	Sequence 97, Appl
9	197.5	10.6	1427	4	US-09-874-923-97	Sequence 97, Appl
10	197.5	10.6	1641	4	US-09-551-974A-96	Sequence 96, Appl
11	197.5	10.6	1641	4	US-09-565-501A-96	Sequence 96, Appl
12	197.5	10.6	1641	4	US-09-639-206A-96	Sequence 96, Appl
13	197.5	10.6	1641	4	US-09-874-923-96	Sequence 96, Appl
14	192	10.3	494	4	US-09-517-779-2	Sequence 2, Appli
15	192	10.3	499	4	US-09-949-016-7370	Sequence 7370, Ap
16	181.5	9.8	563	4	US-09-949-016-7419	Sequence 7419, Ap
17	179.5	9.6	982	4	US-09-551-974A-95	Sequence 95, Appl
18	179.5	9.6	982	4	US-09-565-501A-95	Sequence 95, Appl
19	179.5	9.6	982	4	US-09-639-206A-95	Sequence 95, Appl
20	179.5	9.6	982	4	US-09-874-923-95	Sequence 95, Appl
21	178	9.6	146	3	US-09-227-357-396	Sequence 396, App
22	174	9.3	332	4	US-09-949-016-7518	Sequence 7518, Ap
23	174	9.3	382	3	US-09-301-978C-2	Sequence 2, Appli
24	174	9.3	546	2	US-08-533-669A-2	Sequence 2, Appli
25	174	9.3	546	3	US-09-183-861-2	Sequence 2, Appli
26	174	9.3	546	3	US-09-022-765-2	Sequence 2, Appli
27	174	9.3	546	4	US-09-551-974A-2	Sequence 2, Appli
28	174	9.3	546	4	US-09-565-501A-2	Sequence 2, Appli
29	174	9.3	546	4	US-09-639-206A-2	Sequence 2, Appli
30	174	9.3	546	4	US-09-874-923-2	Sequence 2, Appli
31	174	9.3	546	4	US-08-798-841-2	Sequence 2, Appli
32	164	8.8	144	3	US-09-045-973-9	Sequence 9, Appli
33	163.5	8.8	346	4	US-09-538-092-723	Sequence 723, App
34	158	8.5	508	4	US-09-639-207-1	Sequence 1, Appli
35	153	8.2	494	4	US-09-270-767-43457	Sequence 43457, A
36	151	8.1	124	4	US-09-513-999C-8104	Sequence 8104, Ap
37	149.5	8.0	459	4	US-09-248-796A-14438	Sequence 14438, A
38	143.5	7.7	391	4	US-09-248-796A-16378	Sequence 16378, A
39	140	7.5	303	3	US-09-045-973-1	Sequence 1, Appli
40	140	7.5	668	4	US-09-949-016-11278	Sequence 11278, A
41	139.5	7.5	484	2	US-08-879-260-4	Sequence 4, Appli
42	139.5	7.5	484	3	US-09-231-529-4	Sequence 4, Appli
43	139.5	7.5	484	3	US-08-977-816-4	Sequence 4, Appli
44	139.5	7.5	484	4	US-09-639-207-5	Sequence 5, Appli
45	137	7.4	369	4	US-09-538-092-1166	Sequence 1166, Ap

ALIGNMENTS

RESULT 1
US-09-227-357-393
; Sequence 393, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18

; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 393
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-393

Query Match 24.9%; Score 464; DB 3; Length 196;
Best Local Similarity 46.0%; Pred. No. 6.6e-37;
Matches 92; Conservative 41; Mismatches 61; Indels 6; Gaps 2;

```
Qy        161 KYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIP 220
              | ::|:| : ::||:|:  | |  : | ::| |: ||  :::|  | |: :  |   |||
Db          2 KIKYDWYQTESQVVITLMIKNVQKNDVNVEFSEKELSALVKLPSGEDYNLKLELLHPIIP 61

Qy        221 EKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKA-KK 279
              |:  ::|||||:||:| | | | |  |:  |   |  | ::       :  ||||   :
Db         62 EQSTFKVLSTKIEIKLKKPEAVRWEKLEGQGDVPTPKQFVAD-----VKNLYPSSSPYTR 116

Qy        280 DWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWK 339
              :|||| |:|:||| |||:|||||: |: || |  ::::||| |||:|| ||||||||
Db        117 NWDKLVGEIKEEEKNEKLEGDAALNRLFQQIYSDGSDEVKRAMNKSFMESGGTVLSTNWS 176

Qy        340 DVGAKKVEGSPPDGMELKKW 359
              ||| :||| :||| || ||:
Db        177 DVGKRKVEINPPDDMEWKKY 196
```

RESULT 2
US-09-227-357-401
; Sequence 401, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 401
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-401

Query Match 24.9%; Score 464; DB 3; Length 228;
Best Local Similarity 46.0%; Pred. No. 8.3e-37;
Matches 92; Conservative 41; Mismatches 61; Indels 6; Gaps 2;

```
Qy        161 KYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIP 220
              | ::|:| : ::||:|:  | |  : | ::| |: ||  :::|  | |: :  |   |||
Db         34 KIKYDWYQTESQVVITLMIKNVQKNDVNVEFSEKELSALVKLPSGEDYNLKLELLHPIIP 93
```

```
Qy        221 EKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKA-KK 279
              |:  ::|||||:||:| | | | |  |:  |   |  | ::       :  ||||    :
Db         94 EQSTFKVLSTKIEIKLKKPEAVRWEKLEGQGDVPTPKQFVAD-----VKNLYPSSSPYTR 148

Qy        280 DWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWK 339
              :|||| |:|:||| |||:|||||: |: || |  ::::||| |||:|| ||||||||
Db        149 NWDKLVGEIKEEEKNEKLEGDAALNRLFQQIYSDGSDEVKRAMNKSFMESGGTVLSTNWS 208

Qy        340 DVGAKKVEGSPPDGMELKKW 359
              ||| :||| :||| || ||:
Db        209 DVGKRKVEINPPDDMEWKKY 228


RESULT 3
US-09-270-767-44320
; Sequence 44320, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44320
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-44320

  Query Match             17.4%;  Score 323;  DB 4;  Length 181;
  Best Local Similarity   34.8%;  Pred. No. 2.7e-23;
  Matches   71;  Conservative   39;  Mismatches   68;  Indels   26;  Gaps    5;

Qy        158 SKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSK 217
              ||   |||:| | |:||:|:  |     :  ::  :: : ::      : |    :|
Db          2 SKMSVRHDWYQSETKVVITVLLKNAVDKNYAVEITQKRVHMT-----ADGYELDLKLLHP 56

Qy        218 IIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKA 277
              |: |:  |:  |||||| |||   : |  |           ::    |   |:
Db         57 IVVERSSYKAFSTKVEITLAKETGIRWENL----------EEAIVAAPVKPK-------- 98

Qy        278 KKDWDKLEAEVKK-EEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLST 336
               |:||:| :| :| :||| |  |:|||   |: ||  :  ::::|| ||| || ||||||
Db         99 AKNWDQLVSEEEKIDEKEAK--GEAALTNLFKKIYSSSSPEVQKAMNKSFSESGGTVLST 156

Qy        337 NWKDVGAKKVEGSPPDGMELKKWE 360
              || :|| ::|   ||:| | ::||
Db        157 NWNEVGKERVTVKPPNGTEFREWE 180


RESULT 4
US-09-248-796A-18350
; Sequence 18350, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18350
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (148),(164)
;   OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-18350

  Query Match             13.1%;  Score 244;  DB 4;  Length 388;
  Best Local Similarity   25.9%;  Pred. No. 3.9e-15;
  Matches  103;  Conservative   46;  Mismatches  121;  Indels  128;  Gaps   18;

Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              ||      :|  :|     |:  |   |||||    |      |  |: |: || |   :| ||
Db         15 MAIEQFITKGDDALKSKDYLGAISYYSQAIKENPQAFSPYLKRSTAYQKLKNNDKAKADI 74

Qy         61 NKAIELDPMMHK------AYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLK-E 113
              : |  :      |      |:: |     |  ||  |  | |         ::| : :: :
Db         75 SSAFSIATEKGKRVDIGLCYFKLG-----LVYYQEKKVKLSL--------TQFEKAVEYD 121

Qy        114 CDE--------RIAEESSQAPVKNVEA--------------------TVAATIEDK---- 141
              | |        :   :    |  |||                        ::| | :||
Db        122 CKEPTLFMWKAKAEYDLKNHPEWNVEXDKEEDDDIDLVLGVEXEYTQSMANTEKDKSHEP 181

Qy        142 ------------EDFTNMENT--PPVIE-----------------PPSKPKYRHDYYNSA 170
                         |  :| |:|   |  :                    |   |:| |:| |
Db        182 KIVELDANEESSEKKSNQESTSSAPAAQATTQAPKSTNVDVINKIAPLNVKFRDDWYQSN 241

Qy        171 TEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVP----GEEPYHFQPRLFSKIIPEKCKYQ 226
               ||::||:|| |  : : ::|     || |  |      |  |:  | ||::|:| : ||:
Db        242 EEVIITIYAKKVNEEKLKVEFDTN--SVCISFPSAAASEYKYYLDP-LFAEIVPSESKYK 298

Qy        227 VLSTKVEIRLAKAEQVTWTTL---------DYSGRPKTVPQKISTPAETAPRPSYP-SSK 276
              | |||:|| | | :   | |          |   : | |       |:|      || |||
Db        299 VYSTKLEITLKKKDANKWPELEKQAVEGVTDNQDKDKKV-----DPSELV----YPTSSK 349

Qy        277 AKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDA 314
              | :|:                |:| |          || | |
Db        350 KKINWNNF-----------KIDDD--------DITKKA 368


RESULT 5
US-09-248-796A-16433
; Sequence 16433, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16433
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-16433

  Query Match             11.1%;  Score 206;  DB 4;  Length 65;
  Best Local Similarity   64.9%;  Pred. No. 1.3e-12;
  Matches   37;  Conservative    9;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        304 NKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              | ||| |:|| ||| |||||||:|:||||||:|:| :   |: |  ||||||:|||:
Db          8 NDFFRKIFKDVDEDSRRAMMKSYVQSNGTVLTTSWDEAKDKEFEVLPPDGMEVKKWD 64


RESULT 6
US-09-551-974A-97
; Sequence 97, Application US/09551974A
```

```
; Patent No. 6500437
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C5
;  CURRENT APPLICATION NUMBER: US/09/551,974A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 97
;   LENGTH: 1427
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-551-974A-97

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1427;
  Best Local Similarity   27.3%;  Pred. No. 9.1e-10;
  Matches   71;  Conservative   41;  Mismatches  101;  Indels   47;  Gaps    9;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |     |:   | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|       ||         :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
              |  :  :  :  : : |      :| ||::       | ||| :     |  :::|:|
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
                  :      | : ||: |
Db        779 ANKTLTVEDNGIGMTKADLV 798


RESULT 7
US-09-565-501A-97
; Sequence 97, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Peter Probst
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C6
;  CURRENT APPLICATION NUMBER: US/09/565,501A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  112
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 97
;   LENGTH: 1427
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-565-501A-97

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1427;
  Best Local Similarity   27.3%;  Pred. No. 9.1e-10;
  Matches   71;  Conservative   41;  Mismatches  101;  Indels   47;  Gaps    9;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |     |:   | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|       ||         :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
              |  :  :  :  : : |      :| ||::       | ||| :     |  :::|:|
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
                  :      | : ||: |
Db        779 ANKTLTVEDNGIGMTKADLV 798


RESULT 8
US-09-639-206A-97
; Sequence 97, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C7
;  CURRENT APPLICATION NUMBER: US/09/639,206A
;  CURRENT FILING DATE:  2000-08-14
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
;   LENGTH: 1427
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-639-206A-97

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1427;
  Best Local Similarity   27.3%;  Pred. No. 9.1e-10;
  Matches   71;  Conservative   41;  Mismatches  101;  Indels   47;  Gaps    9;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |     |:   | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|       ||         :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731
```

```
Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
Db        779 ANKTLTVEDNGIGMTKADLV 798


RESULT 9
US-09-874-923-97
; Sequence 97, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Brannon, Mark
;  TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE: 210121.420C8
;  CURRENT APPLICATION NUMBER: US/09/874,923
;  CURRENT FILING DATE:  2001-06-04
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
;   LENGTH: 1427
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-874-923-97

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1427;
  Best Local Similarity   27.3%;  Pred. No. 9.1e-10;
  Matches   71;  Conservative   41;  Mismatches  101;  Indels   47;  Gaps    9;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
Db        779 ANKTLTVEDNGIGMTKADLV 798


RESULT 10
US-09-551-974A-96
; Sequence 96, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C5
;  CURRENT APPLICATION NUMBER: US/09/551,974A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 96
;   LENGTH: 1641
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-551-974A-96

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1641;
  Best Local Similarity   27.3%;  Pred. No. 1.1e-09;
  Matches   71;  Conservative   41;  Mismatches  101;  Indels   47;  Gaps    9;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
Db        779 ANKTLTVEDNGIGMTKADLV 798


RESULT 11
US-09-565-501A-96
; Sequence 96, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Webb, John R.
;  APPLICANT:  Dillion, Davin C.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Coler, Rhea
;  APPLICANT:  Peter Probst
;  TITLE OF INVENTION:  LEISHMANIA ANTIGENS FOR USE IN THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF LEISHMANIASIS
;  FILE REFERENCE:  210121.420C6
;  CURRENT APPLICATION NUMBER: US/09/565,501A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  112
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 96
;   LENGTH: 1641
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
;   OTHER INFORMATION: Leishmania antigens
US-09-565-501A-96

  Query Match             10.6%;  Score 197.5;  DB 4;  Length 1641;
```

Best Local Similarity 27.3%; Pred. No. 1.1e-09;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |      |:    | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|      ||          :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
              |  :  :  :  : : |      :| ||::       | ||| :     |  :::|:|
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
                  :      | : ||: |
Db        779 ANKTLTVEDNGIGMTKADLV 798
```

RESULT 12
US-09-639-206A-96
; Sequence 96, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillion, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-639-206A-96

Query Match 10.6%; Score 197.5; DB 4; Length 1641;
Best Local Similarity 27.3%; Pred. No. 1.1e-09;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |      |:    | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|      ||          :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
              |  :  :  :  : : |      :| ||::       | ||| :     |  :::|:|
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
                  :      | : ||: |
Db        779 ANKTLTVEDNGIGMTKADLV 798
```

RESULT 13
US-09-874-923-96
; Sequence 96, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion (poly-protein) constructs comprising multiple
; OTHER INFORMATION: Leishmania antigens
US-09-874-923-96

Query Match 10.6%; Score 197.5; DB 4; Length 1641;
Best Local Similarity 27.3%; Pred. No. 1.1e-09;
Matches 71; Conservative 41; Mismatches 101; Indels 47; Gaps 9;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |   : :  : | :| |  |   |::||   ||    |::|| |:|||| : :|: || |
Db        566 AKQKKDEGNQYFKEDKFPEAVAAYTEAIKRNPAEHTSYSNRAAAYIKLGAFNDALKDAEK 625

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELG-----SSYASGDSRFARLLKECDER 117
               ||| |   | | ||| |    ::|  |  |   |      |:    | |:  ::|
Db        626 CIELKPDFVKGYARKGHAYFWTKQYNRALQAYNEGLKVDPSNADCKDGRYRTIMK----- 680

Qy        118 IAEESSQAPVKNVEAT--------VAATIED-------KEDFTNMENTPPVIEPPSKPKY 162
              | | :|      ||          :|| ::|       ||    |:| |  |:     :|
Db        681 IQEMASGQSADGDEAARRAMDDPEIAAIMQDSYMQLVLKE----MQNDPTRIQ-----EY 731

Qy        163 RHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEK 222
              |  :  :  :  : : |      :| ||::       | ||| :     |  :::|:|
Db        732 MKD--SGISSKINKLISAG------IIRFGQEFSLTDPAVLGEETH-----LRVRVVPDK 778

Qy        223 CKYQVLSTKVEIRLAKAEQV 242
                  :      | : ||: |
Db        779 ANKTLTVEDNGIGMTKADLV 798
```

RESULT 14
US-09-517-779-2
; Sequence 2, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Xu, Xiang
; TITLE OF INVENTION: Cell Cycle Proteins Associated with Rad9, Compositions
; TITLE OF INVENTION: and Methods Of Use
; FILE REFERENCE: A68293/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/517,779
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 11

```
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 494
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-517-779-2

  Query Match             10.3%;  Score 192;  DB 4;  Length 494;
  Best Local Similarity   24.7%;  Pred. No. 6.2e-10;
  Matches   62;  Conservative   47;  Mismatches   92;  Indels   50;  Gaps    6;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|:::| : |   |:| | : |||||:  |: |  | :|: |:::   |  |: || :
Db         23 AEELKTQANDYFKAKDYENAIKFYSQAIELNPSNAIYYGNRSLAYLRTECYGYALGDATR 82

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||    | |||: |: : | ::: |    |         |       :||:: : :::
Db         83 AIELDKKYIKGYYRRAASNMALGKFRAALRDYETVVKVKPHDKDAKMKYQECNKIVKQKA 142

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182
              :  :  |     :: |  | :|    : :  | ||                       :
Db        143 FERAIAGDEH--KRSVVDSLDIESM----TIEDEYSGPK-------------------L 176

Qy        183 PADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQV----------LSTKV 232
              | | | ::::         : |  | :|       || ||:          ||| |
Db        177 EDGKVTISFMKELM---------QWYKDQKKLH-----RKCAYQILVQVKEVLSKLSTLV 222

Qy        233 EIRLAKAEQVT 243
              |  | : |::|
Db        223 ETTLKETEKIT 233


RESULT 15
US-09-949-016-7370
; Sequence 7370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7370
;   LENGTH: 499
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7370

  Query Match             10.3%;  Score 192;  DB 4;  Length 499;
  Best Local Similarity   24.7%;  Pred. No. 6.3e-10;
  Matches   62;  Conservative   47;  Mismatches   92;  Indels   50;  Gaps    6;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|:::| : |   |:| | : |||||:  |: |  | :|: |:::   |  |: || :
Db         28 AEELKTQANDYFKAKDYENAIKFYSQAIELNPSNAIYYGNRSLAYLRTECYGYALGDATR 87

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||    | |||: |: : | ::: |    |         |       :||:: : :::
Db         88 AIELDKKYIKGYYRRAASNMALGKFRAALRDYETVVKVKPHDKDAKMKYQECNKIVKQKA 147

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182
              :  :  |     :: |  | :|    : :  | ||                       :
Db        148 FERAIAGDEH--KRSVVDSLDIESM----TIEDEYSGPK-------------------L 181

Qy        183 PADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQV----------LSTKV 232
              | | | ::::         : |  | :|       || ||:          ||| |
Db        182 EDGKVTISFMKELM---------QWYKDQKKLH-----RKCAYQILVQVKEVLSKLSTLV 227

Qy        233 EIRLAKAEQVT 243
              |  | : |::|
Db        228 ETTLKETEKIT 238


Search completed: April 20, 2005, 18:35:55
Job time : 45 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:17:15 ; Search time 164 Seconds
(without alignments)
851.344 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1589.5	85.4	367	8	ADQ37151	Adq37151 OsSGT1 pr
2	1589.5	85.4	367	8	ADQ15735	Adq15735 Rice stre
3	1135	61.0	358	3	AAG21460	Aag21460 Arabidops
4	1135	61.0	358	3	AAG52816	Aag52816 Arabidops
5	1134	60.9	350	3	AAG29759	Aag29759 Arabidops
6	1134	60.9	371	3	AAG29758	Aag29758 Arabidops
7	1128	60.6	358	3	AAG24735	Aag24735 Arabidops
8	870	46.7	278	3	AAG52817	Aag52817 Arabidops
9	870	46.7	278	3	AAG21461	Aag21461 Arabidops
10	870	46.7	278	3	AAG24736	Aag24736 Arabidops
11	857	46.1	270	3	AAG29760	Aag29760 Arabidops
12	791	42.5	250	3	AAG24737	Aag24737 Arabidops
13	791	42.5	250	3	AAG21462	Aag21462 Arabidops
14	791	42.5	250	3	AAG52818	Aag52818 Arabidops
15	615	33.0	333	6	ABO07188	Abo07188 Human p53
16	615	33.0	333	8	ADJ66517	Adj66517 Suppresso
17	615	33.0	344	5	ABP41961	Abp41961 Human ova
18	594	31.9	365	5	ABB09748	Abb09748 Amino aci
19	594	31.9	365	6	ABP58341	Abp58341 Human cel
20	508.5	27.3	110	7	ABM74423	Abm74423 DNA clone
21	489	26.3	252	3	AAY73366	Aay73366 HTRM clon
22	464	24.9	196	2	AAY02827	Aay02827 Fragment
23	464	24.9	196	7	ADA07572	Ada07572 Human sec
24	464	24.9	196	8	ADN41381	Adn41381 Novel hum
25	464	24.9	228	7	ADA07580	Ada07580 Human sec
26	464	24.9	228	8	ADN41389	Adn41389 Novel hum
27	445	23.9	413	5	ABP73229	Abp73229 Candida a
28	405	21.8	400	6	ABJ25759	Abj25759 Aspergill
29	383	20.6	478	6	ABJ26359	Abj26359 Aspergill
30	317	17.0	178	4	ABB64536	Abb64536 Drosophil
31	311.5	16.7	283	4	AAM90770	Aam90770 Human imm
32	300	16.1	368	7	ADF75113	Adf75113 A_gossypi
33	224	12.0	482	5	AAE16244	Aae16244 Tomato in
34	224	12.0	485	5	AAE16236	Aae16236 Tomato in
35	203.5	10.9	955	7	ADB78890	Adb78890 Leishmani
36	200.5	10.8	369	3	AAG05677	Aag05677 Arabidops
37	199.5	10.7	714	4	ABG09988	Abg09988 Novel hum
38	197.5	10.6	1427	5	AAU71857	Aau71857 Leishmani
39	197.5	10.6	1427	5	AAB71311	Aab71311 MAPS1A-M1
40	197.5	10.6	1427	7	ADB78860	Adb78860 Leishmani
41	197.5	10.6	1641	5	AAU71856	Aau71856 Leishmani
42	197.5	10.6	1641	5	AAB71310	Aab71310 MAPS1A-M1
43	197.5	10.6	1641	7	ADB78859	Adb78859 Leishmani
44	192	10.3	494	5	ABG70122	Abg70122 Human pre
45	192	10.3	499	4	AAE09732	Aae09732 Protein p

ALIGNMENTS

RESULT 1
ADQ37151
ID ADQ37151 standard; protein; 367 AA.
XX
AC ADQ37151;
XX
DT 07-OCT-2004 (first entry)
XX
DE OsSGT1 protein.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response.
XX
OS Oryza sativa.
XX
PN WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
PS Claim 1; SEQ ID NO 310; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related polypeptide. The
CC present sequence is published separately from the main body of the
CC specification as EPO data.
XX
SQ Sequence 367 AA;

Query Match 85.4%; Score 1589.5; DB 8; Length 367;

```
  Best Local Similarity   83.8%;  Pred. No. 1.7e-137;
  Matches  306;  Conservative   28;  Mismatches   24;  Indels    7;  Gaps    2;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||| |||||||||| |||:|||:| ||||:|||||||||||||||||||||||
Db          5 AASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYADRAQAHIKLGNYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||||||:||||||||||||| |:||||||| ||:||||||||||
Db         65 KAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYSFASGDSRFTRLMKECDERIAEE 124

Qy        122 SSQAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLT 176
               :: |||  |      :||: :|:|:|  ||:||||::|   |||||||:||||||||||
Db        125 LTEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              ||||||||::||:||||||||||||||||||||||||||||| :||||||||||||||||
Db        183 IFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKSRYQVLSTKVEIRL 242

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEK 296
              |||||:|||:|||  :|| |||||  |||:| ||||||||:|||||||||||||||||||
Db        243 AKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRPSYPSSKSKKDWDKLEAEVKKEEKEEK 302

Qy        297 LDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEL 356
              |:|||||||||||| ||||||||||||||||||||||||||||:|||||||||||||||
Db        303 LEGDAALNKFFRDIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVEGSPPDGMEL 362

Qy        357 KKWEY 361
              |||||
Db        363 KKWEY 367


RESULT 2
ADQ15735
ID   ADQ15735 standard; protein; 367 AA.
XX
AC   ADQ15735;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Rice stress-related protein #72.
XX
KW   rice; stress-related protein; plant maturation; plant development;
KW   plant proliferation; plant senescence; plant disease-resistance;
KW   plant stress response; transgenic plant; pest tolerance;
KW   herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
KW   improved nutritional value; increased yield; increased proliferation.
XX
OS   Oryza sativa.
XX
PN   WO2004061080-A2.
XX
PD   22-JUL-2004.
XX
PF   23-DEC-2003; 2003WO-US041098.
XX
PR   26-DEC-2002; 2002US-0436564P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Cooper B;
XX
DR   WPI; 2004-534374/51.
DR   N-PSDB; ADQ15734.
XX
PT   New isolated nucleic acids and proteins, useful for producing transgenic
PT   plants having improved properties, e.g. tolerance to pests, herbicides,
PT   or biotic or abiotic stresses, improved nutritional value, or increased
PT   yield or proliferation.
XX
PS   Claim 27; SEQ ID NO 144; 551pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of rice
CC   stress-related proteins. The DNA and protein sequences of the invention
CC   are useful for regulating and controlling plant maturation and
CC   development, including proliferation, senescence, disease-resistance, or
CC   stress response. They are also useful for producing transgenic plants
CC   having improved properties, e.g. tolerance to pests, herbicides, or
CC   biotic or abiotic stresses, improved nutritional value, increased yield
CC   or proliferation, or improved structure causing less loss from lodging or
CC   shattering. The present amino acid sequence represents a rice stress-
CC   related protein of the invention.
XX
SQ   Sequence 367 AA;

  Query Match             85.4%;  Score 1589.5;  DB 8;  Length 367;
  Best Local Similarity   83.8%;  Pred. No. 1.7e-137;
  Matches  306;  Conservative   28;  Mismatches   24;  Indels    7;  Gaps    2;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||| |||||||||| |||:|||:| ||||:|||||||||||||||||||||||
Db          5 AASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYADRAQAHIKLGNYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||||||:||||||||||||| |:||||||| ||:||||||||||
Db         65 KAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYSFASGDSRFTRLMKECDERIAEE 124

Qy        122 SSQAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLT 176
               :: |||  |      :||: :|:|:|  ||:||||::|   |||||||:||||||||||
Db        125 LTEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              ||||||||::||:||||||||||||||||||||||||||||| :||||||||||||||||
Db        183 IFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKSRYQVLSTKVEIRL 242

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEK 296
              |||||:|||:|||  :|| |||||  |||:| ||||||||:|||||||||||||||||||
Db        243 AKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRPSYPSSKSKKDWDKLEAEVKKEEKEEK 302

Qy        297 LDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEL 356
              |:|||||||||||| ||||||||||||||||||||||||||||:|||||||||||||||
Db        303 LEGDAALNKFFRDIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVEGSPPDGMEL 362

Qy        357 KKWEY 361
              |||||
Db        363 KKWEY 367


RESULT 3
AAG21460
ID   AAG21460 standard; protein; 358 AA.
XX
AC   AAG21460;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 24024.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
```

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.

PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 61.0%; Score 1135; DB 3; Length 358;
Best Local Similarity 61.0%; Pred. No. 1.3e-95;
Matches 224; Conservative 49; Mismatches 76; Indels 18; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||:||||::| :|||:|||  |  |  :||||||:||: |:|||| ||||
Db          2 AKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQANIKIDNFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||| ||:||||| ||||||| |:| |  : :| ::: ||| |||||
Db         62 AIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEE 121

Qy        123 SQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVV 174
                    : :   |  ||  |          |||  |  |:|| :||::|    | |
Db        122 KDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAV 172

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
              :||||| || ::| ::||||:||| |:| ||| || ||||| |||||||:::|||||||
Db        173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEI 232

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
              |||||| :|| :|:| |: ::|  | :  :  : || |||||  ||||||||||||:||:
Db        233 RLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKD 291

Qy        295 EKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGM 354
              ||||||||:|||| |||  |||||||| ||| ||||||||||||:|| |||| :|||||
Db        292 EKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGM 351

Qy        355 ELKKWEY 361
              |||||||
Db        352 ELKKWEY 358
```

RESULT 4
AAG52816
ID AAG52816 standard; protein; 358 AA.
XX
AC AAG52816;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67180.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 61.0%; Score 1135; DB 3; Length 358;
Best Local Similarity 61.0%; Pred. No. 1.3e-95;
Matches 224; Conservative 49; Mismatches 76; Indels 18; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||:||||::| :|||:|||  |  |  :||||||:||: |:|||| ||||
Db          2 AKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQANIKIDNFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||| ||:||||| ||||||| |:| |  : :| ::: ||| |||||
Db         62 AIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEE 121

Qy        123 SQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVV 174
                    : :   |  ||  |          |||  |  |:|| :||::|    | |
Db        122 KDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAV 172

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
              :||||| || ::| ::|||:||| |:| ||| || ||||| |||||||:::||||||||
Db        173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEI 232

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
              |||||| :|| :|:| |: ::|  | :  :  : || |||||  ||||||||||||:||:
Db        233 RLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKD 291

Qy        295 EKLDGDAALNKFFRDIYKDADEDMRRAMMKSPVESNGTVLSTNWKDVGAKKVEGSPPDGM 354
              ||||||||:|||| |||  |||||||||| ||| ||||||||||||:|| |||| :|||||
Db        292 EKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGM 351
```

```
Qy        355 ELKKWEY 361
              |||||||
Db        352 ELKKWEY 358


RESULT 5
AAG29759
ID   AAG29759 standard; protein; 350 AA.
XX
AC   AAG29759;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 35462.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
```

```
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match            60.9%;  Score 1134;  DB 3;  Length 350;
  Best Local Similarity  61.4%;  Pred. No. 1.6e-95;
  Matches  221;  Conservative   56;  Mismatches   69;  Indels   14;  Gaps    4;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||||||||::| :|||:|||  |  |: :||||||:||| ::|||||||||
Db          2 AKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQAYIKLESFTEAVADANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||| : ||| ||| ||:|||||:||| ||| |:|    :|:| :|: ||:  | ||
Db         62 AIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEE 121

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAK 180
                  |: | :|: :::             ||| |    |  ||||:||    |||:|:|||
Db        122 KDL-VQPVPSTLPSSV----------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAK 170

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              |:|  :| ||||||:||| ||||||: |: ||||| ||||:||||:|||||:|| ||||:
Db        171 GIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEICLAKAD 230

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
               :|| :|::   |  :| | :  :| : ||:|||||  |||||||||||:||:|||:||
Db        231 IITWASLEHGKGPAVLP-KPNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGD 289

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||:||::|||||||| |||||||||||||||::|| | :| :|||||||||||
Db        290 AALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 349


RESULT 6
AAG29758
ID   AAG29758 standard; protein; 371 AA.
XX
AC   AAG29758;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 35461.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
```

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.9%; Score 1134; DB 3; Length 371;
Best Local Similarity 61.4%; Pred. No. 1.7e-95;
Matches 221; Conservative 56; Mismatches 69; Indels 14; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  |||||||||||::| :|||:|||  |  |: :||||||:||| ::|||||||||
Db         23 AKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQAYIKLESFTEAVADANK 82

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||| : ||| ||| ||:|||||:||| ||| |:|    :|:| :|: ||:  | ||
Db         83 AIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEE 142

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAK 180
                 |: | :|: :::             ||| |    |  ||||:||    |||:|:|||
Db        143 KDL-VQPVPSTLPSSV----------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAK 191

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              |:|  :| ||||||:||| ||||||: |: ||||| ||||:||||:|||||:|| ||||:
Db        192 GIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEICLAKAD 251

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
               :|| :|::   |  :| | :  :| : ||:|||||  ||||||||||||:||:|||:||
Db        252 IITWASLEHGKGPAVLP-KPNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGD 310

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||:||::|||||||| ||||||||||||||::|| | :| :||||||||||
Db        311 AALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 370
```

RESULT 7
AAG24735
ID AAG24735 standard; protein; 358 AA.
XX
AC AAG24735;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28523.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.6%; Score 1128; DB 3; Length 358;
Best Local Similarity 60.8%; Pred. No. 5.7e-95;
Matches 223; Conservative 49; Mismatches 77; Indels 18; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||| :||||::| :|||:|||  |  |  :||||||:||: |:|||| ||||
Db          2 AKELAEKAKEAXLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQANIKIDNFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||| ||:||||| ||||||| |:| |  : :| ::: ||| |||||
Db         62 AIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEE 121

Qy        123 SQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVV 174
                    : :   |  ||  |          |||  |  |:|| :||::|    | |
Db        122 KDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAV 172

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
              :||||| || ::| ::|||||:||| |:| ||| || ||||| |||||||:::|||||||
Db        173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEI 232

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
              |||||| :|| :|:| |: ::| | :  :  : |' |||||  |||||||||||:||:
Db        233 RLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKD 291

Qy        295 EKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGM 354
              ||||||||:|||| ||| |||||||||| ||| ||||||||||||:|| |||| :|||||
Db        292 EKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGM 351

Qy        355 ELKKWEY 361
              |||||||
Db        352 ELKKWEY 358
```

RESULT 8
AAG52817
ID AAG52817 standard; protein; 278 AA.
XX
AC AAG52817;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67181.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 870; DB 3; Length 278;
Best Local Similarity 60.1%; Pred. No. 2.4e-71;
Matches 173; Conservative 37; Mismatches 60; Indels 18; Gaps 4;

```
Qy        82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQ------APVKNVEATVA 135
             :||||| ||||||| |:| |  : :| ::: ||| |||||           : :   |
Db         1 MKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPL 60

Qy       136 ATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGE 193
             ||  |          |||  |  |:|| :||::|    | |:||||| || ::| ::|||
Db        61 ATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAVVTIFAKKVPKENVTVEFGE 111

Qy       194 QMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRP 253
             |:||| |:| ||| || ||||| |||||||:::|||||||||||||| :|| :|:| |:
Db       112 QILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170

Qy       254 KTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKD 313
             ::|  | :  :  : || |||||  |||||||||||:||:||||||||:|||| |||
Db       171 QSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSS 230

Qy       314 ADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWEY 361
             ||||||||| ||| ||||||||||||:|| |||| :|||||||||||
Db       231 ADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGMELKKWEY 278
```

RESULT 9
AAG21461
ID AAG21461 standard; protein; 278 AA.
XX
AC AAG21461;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24025.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.7%; Score 870; DB 3; Length 278;
Best Local Similarity 60.1%; Pred. No. 2.4e-71;
Matches 173; Conservative 37; Mismatches 60; Indels 18; Gaps 4;

```
Qy         82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQ------APVKNVEATVA 135
              :||||| ||||||| |:| |  : :| ::: ||| |||||          : :   |
Db          1 MKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPL 60

Qy        136 ATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGE 193
              ||  |           |||  |  |:|| :||::|    | |:||||| || ::| ::|||
Db         61 ATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAVVTIFAKKVPKENVTVEFGE 111

Qy        194 QMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRP 253
              |:||| |:| ||| || ||||| |||||||:::|||||||||||||| :|| :|:| |:
Db        112 QILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170

Qy        254 KTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKD 313
              ::|  | :  :  : || |||||  |||||||||||:||:||||||||:|||| |||
Db        171 QSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSS 230

Qy        314 ADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWEY 361
              ||||||||| ||| |||||||||||||:|| |||| :|||||||||||
Db        231 ADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGMELKKWEY 278
```

RESULT 10
AAG24736
ID AAG24736 standard; protein; 278 AA.
XX
AC AAG24736;
XX

DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28524.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

```
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             46.7%;  Score 870;  DB 3;  Length 278;
  Best Local Similarity   60.1%;  Pred. No. 2.4e-71;
  Matches  173;  Conservative   37;  Mismatches   60;  Indels   18;  Gaps    4;

Qy         82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQ------APVKNVEATVA 135
              :||||| ||||||| |:| |  : :| ::: ||| |||||         : :   |
Db          1 MKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEEKDLVQPMPPSLPSSSTTPL 60

Qy        136 ATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGE 193
              ||  |           |||  |  |:|| :||::|    | |:||||| || ::| ::|||
Db         61 ATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAVVTIFAKKVPKENVTVEFGE 111

Qy        194 QMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRP 253
              |:||| |:| ||| || ||||| |||||||:::||||||||||||| :|| :|:| |:
Db        112 QILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEIRLAKAEIITWASLEY-GKG 170

Qy        254 KTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKD 313
              ::|  | :  :  : || |||||  ||||||||||||:||:||||||||:|||| |||
Db        171 QSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSS 230

Qy        314 ADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWEY 361
              ||||||||| ||| ||||||||||||:|| |||| :||||||||||||
Db        231 ADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGMELKKWEY 278


RESULT 11
AAG29760
ID   AAG29760 standard; protein; 270 AA.
XX
AC   AAG29760;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 35463.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
```

PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 857; DB 3; Length 270;
Best Local Similarity 59.4%; Pred. No. 3.6e-70;
Matches 167; Conservative 45; Mismatches 55; Indels 14; Gaps 4;

```
Qy         82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIEDK 141
              :|||||:||| ||| |:|    :|:| :|: ||:  | ||     |: | :|: :::
Db          1 MKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEEKDL-VQPVPSTLPSSV--- 56

Qy        142 EDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVS 199
                  ||| |    |  ||||:||    |||:|:||||:|  :| ||||||:|||
Db         57 -------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAKGIPKQNVNIDFGEQILSVV 109

Qy        200 IEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQK 259
```

```
              ||||||: |: ||||| ||||:||||:|||||:|| ||||: :|| :|:: | :| |
Db        110 IEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEICLAKADIITWASLEHGKGPAVLP-K 168

Qy        260 ISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMR 319
               :  :| : ||:|||||  ||||||||||||:||:|||:||||||||||:||::||||||
Db        169 PNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGDAALNKFFREIYQNADEDMR 228

Qy        320 RAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||| ||||||||||||||::|| | :| :||||||||||
Db        229 RAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 269


RESULT 12
AAG24737
ID   AAG24737 standard; protein; 250 AA.
XX
AC   AAG24737;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 28525.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
```

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.5%; Score 791; DB 3; Length 250;
Best Local Similarity 60.4%; Pred. No. 3.9e-64;
Matches 157; Conservative 32; Mismatches 53; Indels 18; Gaps 4;

```
Qy      110 LLKECDERIAEESSQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPK 161
            :: ||| |||||          : :   |  ||  |          |||  |  |:||
Db        1 MIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPM 51

Qy      162 YRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPE 221
            :||::|    | |:|||| || ::| ::||||:||| |:| ||| || ||||| |||||
Db       52 FRHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPE 111

Qy      222 KCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDW 281
            ||:::|||||||||||||| :|| :|:| |: ::|  | :  :  : || |||||  |||
Db      112 KCRFEVLSTKVEIRLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDW 170

Qy      282 DKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDV 341
            |||||||||:||:||||||||:|||| |||  |||||||| ||| ||||||||||||:|
Db      171 DKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEV 230

Qy      342 GAKKVEGSPPDGMELKKWEY 361
            | |||| :|||||||||||
Db      231 GTKKVESTPPDGMELKKWEY 250
```

RESULT 13
AAG21462
ID AAG21462 standard; protein; 250 AA.
XX
AC AAG21462;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24026.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

```
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             42.5%;  Score 791;  DB 3;  Length 250;
  Best Local Similarity   60.4%;  Pred. No. 3.9e-64;
  Matches  157;  Conservative   32;  Mismatches   53;  Indels   18;  Gaps    4;

Qy        110 LLKECDERIAEESSQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPK 161
              :: ||| |||||          : :   |  ||  |          |||  |  |:||
Db          1 MIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPM 51

Qy        162 YRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPE 221
              :||::|    | |:||||| || ::| ::||||:||| |:| ||| || ||||| |||||
Db         52 FRHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPE 111

Qy        222 KCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDW 281
              ||:::|||||||||||||| :|| :|:| |: ::|  | :  :  : || |||||  |||
Db        112 KCRFEVLSTKVEIRLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDW 170

Qy        282 DKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDV 341
              |||||||||:||:|||||||:|||| |||  ||||||||| ||| ||||||||||||:|
Db        171 DKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEV 230

Qy        342 GAKKVEGSPPDGMELKKWEY 361
              | |||| :|||||||||||
Db        231 GTKKVESTPPDGMELKKWEY 250


RESULT 14
AAG52818
ID   AAG52818 standard; protein; 250 AA.
XX
AC   AAG52818;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 67182.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
```

PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.5%; Score 791; DB 3; Length 250;
Best Local Similarity 60.4%; Pred. No. 3.9e-64;
Matches 157; Conservative 32; Mismatches 53; Indels 18; Gaps 4;

```
Qy        110 LLKECDERIAEESSQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPK 161
              :: ||| |||||           : :   |  ||  |          |||  |  |:||
Db          1 MIDECDLRIAEEEKDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPM 51

Qy        162 YRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPE 221
              :||::|    | |:||||| || ::| ::||||:||| |:| ||| || ||||| |||||
Db         52 FRHEFYQKPEEAVVTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPE 111

Qy        222 KCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDW 281
              ||:::|||||||||||||| :|| :|:| |: ::|  | :  :  : || |||||  |||
Db        112 KCRFEVLSTKVEIRLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDW 170

Qy        282 DKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDV 341
              |||||||||:||:||||||||:|||| |||  ||||||||| ||| |||||||||||||:|
Db        171 DKLEAEVKKQEKDEKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEV 230

Qy        342 GAKKVEGSPPDGMELKKWEY 361
              | |||| :||||||||||||
Db        231 GTKKVESTPPDGMELKKWEY 250
```

RESULT 15
ABO07188
ID ABO07188 standard; protein; 333 AA.
XX
AC ABO07188;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 148.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13363.
XX

PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 456; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 333 AA;

Query Match 33.0%; Score 615; DB 6; Length 333;
Best Local Similarity 38.7%; Pred. No. 9.9e-48;
Matches 135; Conservative 62; Mismatches 116; Indels 36; Gaps 5;

```
Qy          12 EAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMH 71
               :| :|:| : | |  ::|::  |  |  |  ||  || ||||  ||||| |::||:|
Db          20 DALIDEDPQAALEELTKALEQKPDDAQYYCQRAYCHILLGNYCVAVADAKKSLELNPNNS 79

Qy          72 KAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVE 131
                 |  |||      : |  |       |    | |: |:  :| | |  |:  |:: |
Db          80 TAMLRKGICEYHEKNYAAALETFTEGQKLDSADANFSVWIKRCQE--AQNGSESEV---- 133

Qy         132 ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDF 191
                        :|:            : | ::|:| : ::||:|:  | |  : | ::|
Db         134 -----------WTH-----------QSKIKYDWYQTESQVVITLMIKNVQKNDVNVEF 169

Qy         192 GEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSG 251
                |: ||  :::|  | |: :  |   ||||:  ::|||||:||:| | | | |  |:  |
Db         170 SEKELSALVKLPSGEDYNLKLELLHPIIPEQSTFKVLSTKIEIKLKKPEAVRWEKLEGQG 229

Qy         252 RPKTVPQKISTPAETAPRPSYPSSKA-KKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDI 310
                  | | ::       :  ||||    ::|||  |:|:||| |||:||||||: |: |
Db         230 DVPTPKQFVAD-----VKNLYPSSSPYTRNWDKLVGEIKEEEKNEKLEGDAALNRLFQQI 284

Qy         311 YKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 359
               | |  ::::||| |||:|| |||||||| ||| :||| :||| || ||:
Db         285 YSDGSDEVKRAMNKSFMESGGTVLSTNWSDVGKRKVEINPPDDMEWKKY 333
```

Search completed: April 20, 2005, 18:31:17
Job time : 169 secs

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:18:25 ; Search time 40 Seconds
(without alignments)
868.356 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1135	61.0	358	2	T13017	hypothetical prote
2	1134	60.9	350	2	T05589	hypothetical prote
3	383	20.6	198	2	T20305	hypothetical prote
4	364.5	19.6	395	2	S66940	SGT1 protein - yea
5	341	18.3	444	2	T40307	hypothetical prote
6	327	17.6	469	2	T48783	related to SGT1 pr
7	223	12.0	533	2	E84858	phosphoprotein pho
8	193.5	10.4	572	2	H86257	protein F5O11.2 [i
9	192	10.3	498	1	S52570	phosphoprotein pho
10	187	10.0	358	2	T00954	hypothetical prote
11	186	10.0	499	1	A55346	phosphoprotein pho
12	181.5	9.8	543	2	A38093	transformation-sen
13	180.5	9.7	473	1	T40391	phosphoprotein pho
14	177.5	9.5	589	2	A32567	stress-induced pro
15	176	9.5	569	2	S56658	stress-induced pro
16	174	9.3	349	2	T08782	hypothetical prote
17	174	9.3	526	1	T45058	phosphoprotein pho
18	169.5	9.1	513	1	S52571	phosphoprotein pho
19	163.5	8.8	346	2	S61991	hypothetical prote
20	163.5	8.8	591	2	T41531	activator of Hsp70
21	161.5	8.7	558	2	T48150	stress-induced pro
22	160	8.6	627	2	T04562	hypothetical prote
23	159.5	8.6	479	1	T46576	phosphoprotein pho
24	159.5	8.6	591	2	T51996	hypothetical prote
25	157.5	8.5	901	2	JC7111	tetratricopeptide
26	151	8.1	677	2	T45682	hypothetical prote
27	148.5	8.0	358	2	T37805	hypothetical prote
28	147	7.9	317	2	T37851	hypothetical prote
29	146.5	7.9	385	2	S40699	hypothetical prote
30	145.5	7.8	422	2	T24865	hypothetical prote
31	141.5	7.6	1979	2	JW0059	mtprd protein - mo
32	139	7.5	605	2	AF1917	hypothetical prote
33	139	7.5	781	2	F86457	unknown protein, 3
34	137.5	7.4	547	2	AE1884	hypothetical prote
35	136.5	7.3	222	2	F86424	unknown protein, 6
36	136	7.3	225	2	AE2539	hypothetical prote
37	132.5	7.1	320	2	T03899	hypothetical prote
38	130.5	7.0	297	2	T34141	hypothetical prote
39	130.5	7.0	1621	2	A82255	hypothetical prote
40	130	7.0	619	2	A36682	72K mitochondrial
41	130	7.0	2025	2	JC5020	tetratricopeptide
42	130	7.0	2055	2	T31110	extracellular matr
43	129.5	7.0	427	2	T00960	hypothetical prote
44	129.5	7.0	542	2	AD1333	class I heat-shock
45	127.5	6.9	961	2	T32493	unc-45 protein - C

ALIGNMENTS

RESULT 1
T13017
hypothetical protein F8L21.50 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13017
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mev
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17587
A;Accession: T13017
A;Molecule type: DNA
A;Residues: 1-358 <BEV>
A;Cross-references: UNIPROT:Q9SUT5; EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.50
A;Experimental source: cultivar Columbia; BAC clone F8L21
C;Genetics:
A;Gene: ATSP:F8L21.50
A;Map position: 4
A;Introns: 54/1; 77/2; 119/1; 161/2; 193/3; 215/3; 262/1; 288/3; 323/3

Query Match 61.0%; Score 1135; DB 2; Length 358;
Best Local Similarity 61.0%; Pred. No. 1.5e-69;
Matches 224; Conservative 49; Mismatches 76; Indels 18; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||:||||::| :|||:|||  |  |  :||||||:||: |:|||| ||||
Db          2 AKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQANIKIDNFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||| ||:||||| ||||||| |:| |  : :| ::: ||| |||||
Db         62 AIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEE 121

Qy        123 SQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVV 174
                    : :  |  ||  |          |||  |  |:|| :||::|    | |
Db        122 KDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAV 172

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
              :||||| || ::| ::||||:||| |:| ||| || ||||| |||||||:::||||||||
Db        173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEI 232

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
              |||||| :|| :|:| |: ::|  | :  :  : || |||||  ||||||||||||:||:
Db        233 RLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKD 291

Qy        295 EKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGM 354
              ||||||||:|||| |||  |||||||||| ||| ||||||||||||:|| |||| :|||||
Db        292 EKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGM 351

Qy        355 ELKKWEY 361
              |||||||
Db        352 ELKKWEY 358
```

RESULT 2
T05589
hypothetical protein F9D16.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05589
R;Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15419
A;Accession: T05589
A;Molecule type: DNA
A;Residues: 1-350 <BEV>
A;Cross-references: UNIPROT:Q9SUR9; EMBL:AL035394
A;Experimental source: cultivar Columbia; BAC clone F9D16
C;Genetics:
A;Map position: 4
A;Introns: 54/1; 77/2; 119/1; 153/2; 185/3; 207/3; 254/1; 280/3; 315/3
A;Note: F9D16.40

Query Match 60.9%; Score 1134; DB 2; Length 350;
Best Local Similarity 61.4%; Pred. No. 1.7e-69;
Matches 221; Conservative 56; Mismatches 69; Indels 14; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||||||||::| :|||:|||  |  |: :||||||:||| ::|||||||||
Db          2 AKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQAYIKLESFTEAVADANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||| : ||| ||| ||:|||||:||| ||| |:|    :|:| :|: ||:  | ||
Db         62 AIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEE 121

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAK 180
                |: | :|: :::           ||| |    |  ||||:||    |||:|:|||
Db        122 KDL-VQPVPSTLPSSV----------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAK 170

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              |:|  :| ||||||:||| ||||||: |: ||||| ||||:||||:|||||:|| ||||:
Db        171 GIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEICLAKAD 230

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
               :|| :|::   |  :| | :  :| : ||:|||||  ||||||||||||:||:|||:||
Db        231 IITWASLEHGKGPAVLP-KPNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGD 289

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||:||::|||||||||| |||||||||||||||::|| | :| :||||||||||
Db        290 AALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 349
```

RESULT 3
T20305
hypothetical protein D1054.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20305
R;Matthews, P.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19255
A;Accession: T20305
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-198 <WIL>
A;Cross-references: UNIPROT:Q18949; EMBL:Z74030; PIDN:CAA98442.1; GSPDB:GN00023; CESP:D1
A;Experimental source: clone D1054
C;Genetics:
A;Gene: CESP:D1054.3
A;Map position: 5
A;Introns: 6/2; 71/3; 164/2

Query Match 20.6%; Score 383; DB 2; Length 198;
Best Local Similarity 41.7%; Pred. No. 5.8e-19;
Matches 86; Conservative 38; Mismatches 68; Indels 14; Gaps 6;

```
Qy        158 SKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSK 217
              |||  |||:: | |:| |||  :||| |   :   :   : :: |   :   |  :| :
Db          3 SKP--RHDWFQSETDVTLTILKRGVPLDDCSVSLSD---NNTLTVKQCDEILFYGQLSGQ 57

Qy        218 IIPEKCKYQVLSTKVEIRLAK-AEQVTWTTL--DYSGRPKTVPQKISTPAETAPRPSYPS 274
              :  :    :  : |||:|| | |    | :|  |  |   :| | :|   |:|     |:
Db         58 VKKDDLTVKCTAAKVEVRLPKFARNERWASLLKDGQGVAASV-QSVSPNPESA-----PT 111

Qy        275 SKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVL 334
              :  ||:|| :| :  |||::| |:||||:|| || :| || :|:||||||: |||||||
Db        112 TTVKKNWDAIEKQAVKEEEDESLEGDAAVNKMFRKMYNDASDDVRRAMMKSYSESNGTVL 171

Qy        335 STNWKDVGAKKVEGSPPDGMELKKWE 360
              |||| ::| || |  ||  || |::|
Db        172 STNWSEIGQKKTECQPPACMEYKEYE 197
```

RESULT 4
S66940
SGT1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O2805; protein YOR057w
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66940
R;Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66929
A;Accession: S66940
A;Molecule type: DNA
A;Residues: 1-395 <BOH>
A;Cross-references: UNIPROT:Q08446; EMBL:Z74965; NID:g1420194; PID:e251976; PID:g1420195
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SGT1; GSS1; MIPS:YOR057w
A;Cross-references: SGD:S0005583; MIPS:YOR057w
A;Map position: 15R
C;Function:
A;Description: suppressor of G2 allele of SKP1

Query Match 19.6%; Score 364.5; DB 2; Length 395;
Best Local Similarity 29.3%; Pred. No. 2.6e-17;
Matches 108; Conservative 69; Mismatches 133; Indels 59; Gaps 14;

```
Qy         16 DDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN-KAIELDPMMHKAY 74
              |   | | ||  :|:     ||:   ||::  | | |:    |   | |  ||     ::|
Db         61 DATMENAKELLDKAL----MTAEGRGDRSK--IGLVNFRYFVHFFNIKDYEL----AQSY 110

Qy         75 YRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVEATV 134
              ::|             ||    :  :    : |    | : :::  : :::  :| ||
Db        111 FKK-----------AKNLGYVDDTLPLWEDRLETKLNKKNKKQKDSTNKHTIKPVE--- 155

Qy        135 AATIEDKEDFTNMEN--TPPVIE--PPSKPKYRHDYYNSATEVVLTIFAKGVPADSVVID 190
                :||:: |  :  :  :|  ||  |   ||:: |:| |:| | :::|   :|     ::
Db        156 --SIENRGDNNSSHSPISPLKIETAPQESPKFKIDWYQSSTSVTISLFTVNLPESKEQVN 213

Qy        191 F-----GEQMLSVSIEVP--GEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVT 243
                    : ||:| :||  | | : :  :|  :: |:    ::   |:|| |:| :
Db        214 IYISPNDRRTLSISYQVPKSGSE-FQYNAKLSHEVDPKAVSLKIFPKKLEITLSKIDSTQ 272

Qy        244 WTTL-----------DYSGRPKTVPQKISTPAETAPRPSYPSSKAKK-DWDKLEAEVKK 290
              | |           |     :  : :|    :  | |||||  || || ||: : :
Db        273 WKKLEEDILTESSRLSDEGKNSDSATRLLSAETASKERLSYPSSSKKKIDWSKLDIDEEA 332

Qy        291 EEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSP 350
              :|:    |       ||: :|  || | :|||||||:|||||  |||:|:||    |: ||
Db        333 DEEAGSAD------SFFQKLYAGADPDTKRAMMKSFIESNGTALSTDWEDVSKGTVKTSP 386

Qy        351 PDGMELKKW 359
              |:||| | |
Db        387 PEGMEPKHW 395
```

RESULT 5
T40307

```
hypothetical protein SPBC36.12c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40307
R;Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21919
A;Accession: T40307
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-444 <LYN>
A;Cross-references: EMBL:AL023589; PIDN:CAA19060.1; GSPDB:GN00067; SPDB:SPBC36.12c
A;Experimental source: strain 972h-; cosmid c36
C;Genetics:
A;Gene: SPDB:SPBC36.12c
A;Map position: 2
A;Introns: 151/2; 282/3

  Query Match             18.3%;  Score 341;  DB 2;  Length 444;
  Best Local Similarity   31.3%;  Pred. No. 1.2e-15;
  Matches   98;  Conservative   53;  Mismatches  114;  Indels   48;  Gaps   11;

Qy          77 KGAACIKLEEYQTAKAALELGSSYASGDS------RFARLLKECDERIAEESS-----QA 125
               :| |   ||:||::         :   :       :   :|    |:  :::|     :
Db         150 RGLAYYLLEQYQSSAICFGFALKHLKKEDLTKWQVQLDSMLTIVHEQQNQDTSSLIPDEC 209

Qy         126 PVKNVEATVAATIEDKEDFTNMENTPP-----VIEPPSKPKYRHDYYNSATEVVLTIFAK 180
               |  |:    || ||  |:| ||     |       |     : |:|:  ::  : : |:||
Db         210 P--NIPELEAAKIEGDEEFLNMSLKAPEGQIEKNEEKLSNRIRYDWSQTSFSLNIDIYAK 267

Qy         181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQ---PRLFSKIIPEKCKYQVLSTKVEIRL- 236
               |  : | :   :  | : |::  |:   |       |: :|:|||  ::: |:|||| |
Db         268 KVKDEDVSLLMEKNTLKIEIKL--EDGSIFSLVLDPLYEEIVPEKSSFKLFSSKVEITLI 325

Qy         237 AKAEQVTWTTLDYSGRPKTVPQKISTPAETA----PRPSYPSSKA------KKDWDKLEA 286
               |  :: |  |            : :||  :     : |  || :       ||||  |
Db         326 KKVSEIKWEAL------------VKSPANNSVNVYAKDSNHSSASGNTKNKAKDWDSLAK 373

Qy         287 EVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKV 346
                  || |    |:|||   |:::||:||:| |||||||: ||||| |||||||| :|
Db         374 LADLEEDEP--TGEAALANLFQNLYKNADDDTRRAMMKSYTESNGTALSTNWKDVKSKTF 431

Qy         347 EGSPPDGMELKKW 359
               |  || ||| ||:
Db         432 ETKPPQGMEPKKF 444


RESULT 6
T48783
related to SGT1 protein [imported] - Neurospora crassa
N;Alternate names: protein 13E11.310
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: T48783
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48783
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-469 <SCH>
A;Cross-references: UNIPROT:Q9P6Y5; EMBL:AL353820; GSPDB:GN00112; NCSP:13E11.310
A;Experimental source: cosmid contig 13E11; strain 74
C;Genetics:
A;Gene: NCSP:13E11.310
A;Map position: 2
A;Introns: 277/3

  Query Match             17.6%;  Score 327;  DB 2;  Length 469;
  Best Local Similarity   24.1%;  Pred. No. 1.1e-14;
  Matches  117;  Conservative   71;  Mismatches  152;  Indels  146;  Gaps   16;

Qy           1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
                 |:|: :  |  :|    |   |     :|::  :  | | | |::|| |  |   |: ||
Db           1 MSATTIAQKGIDALAKGDHATAITNLDKALESSNSPAWLLA-RSKAHQKAKNLEAALHDA 59

Qy          61 NKAIELDPMMHKAYYRKGAACIKLE----EYQTAKAALELGSSYASGD--SRFARLLKE- 113
                  |       : |   :|:   :      :|: |    :||  :|  |  :::: || |
Db          60 ELA-------YHAAAERGSGTSRSHMIQAQYRRAVIYYQLG-RFADADCCAKWSMLLAEG 111

Qy         114 ----------------------CDERIAEESSQAPVKNVEATVAATIED---------- 140
                                    || :|:: :|   |  :    |  |
Db         112 RPAREDDGVEKKVDSEGNYTVTYDEFLADKENQPKPKTSDGNALAAAEKGPYSTDWNMAF 171

Qy         141 --------------------KEDFTNM------------------------ENTPPVIEP 156
                                   | : | :                        |  ||  |
Db         172 SWRSQALGRLKNLPKDHPGWKVNVTKIPPKPEKKKAKSPEPAESASEDELKEQKPPQKEA 231

Qy         157 PS-------KPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGE-QMLSVSIE-VPGEEP 207
               |:       | | | |:| :   | :::| | |  : : ::||: |   | |  :| |
Db         232 PAPGSVSDEKMKLRIDFYQTNQTVTVSLFVKDVKKEDLKVEFGKRQRTKVRISPIPREAA 291

Qy         208 YHFQPR---------LFSKIIPEKCKYQVLSTKVEIRLAKA------------------ 239
               : :|          |  :| |   ::      |:|: | ||
Db         292 PYVKPGDRQATSTLVLAGEIDPSASRWSASPRKIELVLQKATPGVKWGRWGEEKIGIVES 351

Qy         240 ----EQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSS--KAKKDWDKLEAEVKKEEK 293
                     :| |:   :  |       ||||:    |:||:|     |:|| |  :  ||:
Db         352 DDQEPAITATSSSAASTAKPALPSTSTPAKV---PAYPTSSKSGPKNWDSLPVDDDKEDG 408

Qy         294 EEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDG 353
               ::       :| ||: :|| :  : :|||||||:|||| |||||  :|  | |   ||:|
Db         409 QD-------INGFFKTLYKGSTPEQQRAMMKSFLESNGTTLSTNWDEVKDKVVPTVPPEG 461

Qy         354 MELKKW 359
               :| | |
Db         462 VEPKPW 467


RESULT 7
E84858
phosphoprotein phosphatase (EC 3.1.3.16) At2g42810 [similarity] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 28-Apr-2003
C;Accession: E84858
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <STO>
A;Cross-references: GB:AE002093; NID:g4512673; PIDN:AAD21727.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g42810
A;Map position: 2
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos
C;Keywords: phosphoric monoester hydrolase

  Query Match             12.0%;  Score 223;  DB 2;  Length 533;
  Best Local Similarity   28.0%;  Pred. No. 1.5e-07;
  Matches   74;  Conservative   43;  Mismatches   99;  Indels   48;  Gaps    9;

Qy           3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
               | : :|:| |||     :  | :||::||:     |  :|:|| || ||  |  |: ||:|
Db          13 AEEFKSQANEAFKGHKYSSAIDLYTKAIELNSNNAVYWANRAFAHTKLEEYGSAIQDASK 72

Qy          63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIA--- 119
               |||:|    | |||:||| : : ::: |    :      :  |    | ||||:: :
Db          73 AIEVDSRYSKGYYRRGAAYLAMGKFKDALKDFQQVKRLSPNDPDATRKLKECEKAVMKLK 132

Qy         120 -EESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIF 178
                ||:   ||     :|| :|    ||  : | |     |:|         :    |: :
```

```
Db        133 FEEAISVPVSE-RRSVAESI----DFHTIGNKPRSSSMPTKTALA-----AVVAAVMVVA 182

Qy        179 AKGVPADSVVIDFGEQMLSVSI-------EVPGEEPYHFQPRL--------FSKIIPEKC 223

Db        183 VRGFATTEIL------MVLVSVVLGTFWWEV---EPQYSGARIEGEEVTLDFVKTMMEDF 233

Qy        224 K----------YQVLSTKVEIRLA 237

Db        234 KNQKTLHKRYAYQIVLQTRQILLA 257


RESULT 8
H86257
protein F5O11.2 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86257
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-572 <STO>
A;Cross-references: UNIPROT:Q9LNB6; GB:AE005172; NID:g8778620; PIDN:AAF79628.1; GSPDB:GN
C;Genetics:
A;Gene: F5O11.2
A;Map position: 1

  Query Match             10.4%;  Score 193.5;  DB 2;  Length 572;
  Best Local Similarity   24.1%;  Pred. No. 1.6e-05;
  Matches   82;  Conservative   50;  Mismatches  121;  Indels   87;  Gaps   13;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62

Db          2 AEEAKAKGNAAFSSGDFTTAINHFTEAIALAPTNHVLFSNRSAAHASLHQYAEALSDAKE 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122

Db         62 TIKLKPYWPKGYSRLGAAHLGLNQFELAVTAYKKG----------------LDVDPTNEA 105

Qy        123 SQAPVKNVEATVAAT----------IEDKEDFTNMENTPPV---IEPPSKPKYRHDYYNS 169

Db        106 LKSGLADAEASVARSRAAPNPFGDAFQGPEMWTKLTSDPSTRGFLQQP-------DFVNM 158

Qy        170 ATEV-----VLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCK 224

Db        159 MQEIQKNPSSLNLYLK----DQRVM----QSLGVLLNVKFRPP---PPQGDEAEVPESDM 207

Qy        225 YQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKL 284

Db        208 GQSSSNEPEV----------------------EKKREP-EPEPEPEVTEEKEKKE---- 239

Qy        285 EAEVKKEEKE------EKLDGDAALNKFFRDIYKDADEDM 318

Db        240 RKEKAKKEKELGNAAYKKKDFETAIQHYSTAIEID-DEDI 278


RESULT 9
S52570
phosphoprotein phosphatase (EC 3.1.3.16) 5 [validated] - human
N;Alternate names: serine/threonine phosphatase PP5
C;Species: Homo sapiens (man)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: S52570; PC4136
R;Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.
EMBO J. 13, 4278-4290, 1994
A;Title: A novel human protein serine/threonine phosphatase, which possesses four tetrat
A;Reference number: S52570; MUID:95009929; PMID:7925273
A;Accession: S52570
A;Molecule type: mRNA
A;Residues: 7-498 <CH2>
A;Cross-references: UNIPROT:P53041; EMBL:S73586; EMBL:X89416
A;Experimental source: teratocarcinoma cell line NTERA-2
R;Xu, X.L.; Lagercrantz, J.; Zickert, P.; Bajalica-Lagercrantz, S.; Zetterberg, A.
Biochem. Biophys. Res. Commun. 218, 514-517, 1996
A;Title: Chromosomal localization and 5' sequence of the human protein serine/threonine
A;Reference number: PC4136; MUID:96144708; PMID:8561788
A;Accession: PC4136
A;Molecule type: mRNA
A;Residues: 1-37 <XUX>
A;Cross-references: EMBL:X92121; NID:g1177477; PIDN:CAA63089.1; PID:e205526; PID:g117747
A;Experimental source: fetal brain
C;Comment: mRNA encoding this protein was detected in all human tissues tested and was l
C;Genetics:
A;Gene: GDB:PPP5C; PPP5; PP5
A;Cross-references: GDB:136857
A;Map position: 19q13.3-19q13.3
C;Function:
A;Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine [vali
A;Note: may play a role in the regulation of RNA synthesis and mitosis
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos
C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F;28-61/Domain: tetratricopeptide repeat homology <TT1>
F;62-95/Domain: tetratricopeptide repeat homology <TT2>
F;96-129/Domain: tetratricopeptide repeat homology <TT3>
F;204-467/Domain: phosphoprotein phosphatase homology <PPP>
F;236-305/Domain: phosphoesterase core homology <PEC>
F;242,244,271/Binding site: iron (Asp, His, Asp) #status predicted
F;271,303,352,426/Binding site: zinc (Asp, Asn, His, His) #status predicted
F;274,304,450/Active site: Asp, His, Tyr #status predicted
F;275,399/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             10.3%;  Score 192;  DB 1;  Length 498;
  Best Local Similarity   24.7%;  Pred. No. 1.7e-05;
  Matches   62;  Conservative   47;  Mismatches   92;  Indels   50;  Gaps    6;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62

Db         28 AEELKTQANDYFKAKDYENAIKFYSQAIELNPSNAIYYGNRSLAYLRTECYGYALGDATR 87

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122

Db         88 AIELDKKYIKGYYRRAASNMALGKFRAALRDYETVVKVKPHDKDAKMKYQECNKIVKQKA 147

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182

Db        148 FERAIAGDEH--KRSVVDSLDIESM----TIEDEYSGPK--------------------L 181

Qy        183 PADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQV----------LSTKV 232

Db        182 EDGKVTISFMKELM---------QWYKDQKKLH-----RKCAYQILVQVKEVLSKLSTLV 227

Qy        233 EIRLAKAEQVT 243

Db        228 ETTLKETEKIT 238


RESULT 10
T00954
hypothetical protein F20D22.4 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00954
R;Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji, O
 K.; Feng, J.; Kim, C.; Kurtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.;
submitted to the EMBL Data Library, May 1998
A;Description: Arabidopsis thaliana chromosome 1 BAC F20D22 complete sequence.
```

```
A;Reference number: Z14214
A;Accession: T00954
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358 <VYS>
A;Cross-references: UNIPROT:O64488; EMBL:AC002411; NID:g2570223; PID:g3142292; GSPDB:GN
C;Genetics:
A;Gene: ATSP:F20D22.4
A;Map position: 1
A;Introns: 53/2; 91/3; 108/3; 166/3; 226/3; 266/3; 311/2

  Query Match           10.0%;  Score 187;  DB 2;  Length 358;
  Best Local Similarity   20.6%;  Pred. No. 2.4e-05;
  Matches   76;  Conservative   54;  Mismatches  125;  Indels  114;  Gaps    8;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |    |: |  | |    :|  |  ||:|||    |: | ||::|| | : |   ::|:|||
Db         14 AEKSLKEKGNEFFKAGNFLKAAALYTQAIKLDPSNATLYSNRAAAFLSLVKLSKALADAE 73

Qy         62 KAIELDPMMHK-------AYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKEC 114
                |:|:|   |        |:|||     :|:|: | || |:    |    :  :| :|
Db         74 TTIKLNPQWEKVKSIHDNGYFRKGCVLEAMEKYEDALAAFEMALQYNPQSTEVSRKIKRL 133

Qy        115 DERIAEESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVV 174
               :   |:     ::|: : |
Db        134 GQLQKEKQRAQELENLRSNV---------------------------------------- 153

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
                          |  : : |              |:  :       |:|  :: |  ||
Db        154 ---------------DMAKHLES------------FKSEMSENYGTEECWKEMFSFIVE- 185

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
                        |:: : :      |: |      |  :   | |   ||    |
Db        186 ----------TMETAVKSWHETSKVDT------RVYFLLDKEKTQTDKYAPAVN----- 223

Qy        295 EKLDGDAALNKFFR--DIYKDADEDMRRAMMKSFVESNGTVLSTN-------WKDVGAKK 345
                    ::| |:  | : :    :|:   :|| ::   |   :       || ||::|
Db        224 --------IDKAFQSPDTHSNCFTYLRQYAEESFSKAACLVTPKSSISYPQVWKGVGSRK 275

Qy        346 VEGSPPDGM 354
               :  | ||:
Db        276 WKLGPNDGI 284


RESULT 11
A55346
phosphoprotein phosphatase (EC 3.1.3.16) PPT [validated] - rat
N;Alternate names: serine/threonine phosphatase PP5
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
C;Accession: A55346
R;Becker, W.; Kentrup, H.; Klumpp, S.; Schultz, J.E.; Joost, H.G.
J. Biol. Chem. 269, 22586-22592, 1994
A;Title: Molecular cloning of a protein serine/threonine phosphatase containing a putati
A;Reference number: A55346; MUID:94357899; PMID:8077208
A;Accession: A55346
A;Molecule type: mRNA
A;Residues: 1-499 <BEC>
A;Cross-references: GB:X77237
A;Note: authors translated the codon AAG for residue 53 as Gln, and GTA for residue 496
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos
C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F;28-61/Domain: tetratricopeptide repeat homology <TT1>
F;62-95/Domain: tetratricopeptide repeat homology <TT2>
F;96-129/Domain: tetratricopeptide repeat homology <TT3>
F;204-468/Domain: phosphoprotein phosphatase homology <PPP>
F;236-305/Domain: phosphoesterase core homology <PEC>
F;242,244,271/Binding site: iron (Asp, His, Asp) #status predicted
F;271,303,352,427/Binding site: zinc (Asp, Asn, His, His) #status predicted
F;274,304,451/Active site: Asp, His, Tyr #status predicted
F;275,400/Binding site: substrate phosphate (Arg) #status predicted

  Query Match           10.0%;  Score 186;  DB 1;  Length 499;
  Best Local Similarity   24.3%;  Pred. No. 4.3e-05;
  Matches   61;  Conservative   46;  Mismatches   94;  Indels   50;  Gaps    6;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|:::| : |   |:| | : ||:||:  |: |  | :|: |:::   |  |: || :
Db         28 AEELKTQANDYFKAKDYENAIKFYSKAIELNPSNAIYYGNRSLAYLRTECYGYALGDATR 87

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||    | |||: |: : | ::: |    |          |      :|| : : :::
Db         88 AIELDKKYIKGYYRRAASNMALGKFRAALRDYETVVKVKPNDKDAKMKYQECSKIVKQKA 147

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182
               :  :   |     :: | |  :|     : :  | ||                    :
Db        148 FERAIAGDEH--RRSVVDSLDIESM----TIEDEYSGPK-------------------L 181

Qy        183 PADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQV----------LSTKV 232
                 | | : ::         :  |  | :|       || ||:          ||| |
Db        182 EDGKVTITFMKDLM---------QWYKDQKKLH-----RKCAYQILVQVKEVLCKLSTLV 227

Qy        233 EIRLAKAEQVT 243
              |  | : |::|
Db        228 ETTLKETEKIT 238


RESULT 12
A38093
transformation-sensitive protein IEF SSP 3521 - human
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C;Accession: A38093
R;Honore, B.; Leffers, H.; Madsen, P.; Rasmussen, H.H.; Vandekerckhove, J.; Celis, J.E.
J. Biol. Chem. 267, 8485-8491, 1992
A;Title: Molecular cloning and expression of a transformation-sensitive human protein co
A;Reference number: A38093; MUID:92235077; PMID:1569099
A;Accession: A38093
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-543 <HON>
A;Cross-references: UNIPROT:P31948; GB:M86752; NID:g184564; PIDN:AAA58682.1; PID:g184565
C;Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat hc
F;4-37/Domain: tetratricopeptide repeat homology <TT1>
F;38-71/Domain: tetratricopeptide repeat homology <TT2>
F;225-258/Domain: tetratricopeptide repeat homology <TT3>
F;259-292/Domain: tetratricopeptide repeat homology <TT4>
F;360-393/Domain: tetratricopeptide repeat homology <TT5>
F;394-427/Domain: tetratricopeptide repeat homology <TT6>
F;428-461/Domain: tetratricopeptide repeat homology <TT7>

  Query Match           9.8%;  Score 181.5;  DB 2;  Length 543;
  Best Local Similarity   22.4%;  Pred. No. 9.7e-05;
  Matches   71;  Conservative   53;  Mismatches  118;  Indels   75;  Gaps   10;

Qy          4 SDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKA 63
              ::|: |  :|    : : | : ||:||   |    ||::|: |: | |:| :|  |  |
Db          5 NELKEKGNKALSVGNIDDALQCYSEAIKLDPHNHVLYSNRSAAYAKKGDYQKAYEDGCKT 64

Qy         64 IELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESS 123
              ::| |   | | || ||   |  :: ||   | |  : : : :    |:  : |:||
Db         65 VDLKPDWGKGYSRKAAALEFLNRFEEAKRTYEEGLKHEANNPQLKEGLQNMEARLAERKF 124

Qy        124 QAP--VKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKG 181
                 | : |:          :  :|: |      | | ||        |::  :  :
Db        125 MNPFNMPNL------------YQKLESDPRTRTLLSDPTYR--------ELIEQL--RN 161

Qy        182 VPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQ 241
              |:|          |   ::          ||: :           ||  : : |   ::
Db        162 KPSD----------LGTKLQ---------DPRIMT----------TLSVLLGVDLGSMDE 192

Qy        242 VTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKD--WDKLEAEVKKEEKEEKLDG 299
                           ::|:||    | |  |  : | :   : |    |:  ||::|
Db        193 --------------EEEIATP----PPPPPPKKETKPEPMEEDLPENKKQALKEKELGN 233

Qy        300 DAALNKFFRDIYKDADE 316
```

```
                ||   | |    |  |:
Db        234 DAYKKKDFDTALKHYDK 250
```

RESULT 13
T40391
phosphoprotein phosphatase (EC 3.1.3.16) SPBC3F6.01c [similarity] - fission yeast (Schiz
N;Alternate names: serine/threonine phosphatase PP5 homolog
C;Species: Schizosaccharomyces pombe
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
C;Accession: T40391
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Churcher, C.M.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21925
A;Accession: T40391
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-473 <LYN>
A;Cross-references: EMBL:AL022019; PIDN:CAA17690.2; GSPDB:GN00067; SPDB:SPBC3F6.01c
A;Experimental source: strain 972h-; cosmid c3F6
C;Genetics:
A;Gene: SPDB:SPBC3F6.01c
A;Map position: 2
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos
C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F;5-38/Domain: tetratricopeptide repeat homology <TT1>
F;39-72/Domain: tetratricopeptide repeat homology <TT2>
F;73-106/Domain: tetratricopeptide repeat homology <TT3>
F;179-444/Domain: phosphoprotein phosphatase homology <PPP>
F;211-280/Domain: phosphoesterase core homology <PEC>
F;217,219,246/Binding site: iron (Asp, His, Asp) #status predicted
F;246,278,327,404/Binding site: zinc (Asp, Asn, His, His) #status predicted
F;249,279,428/Active site: Asp, His, Tyr #status predicted
F;250,377/Binding site: substrate phosphate (Arg) #status predicted

  Query Match 9.7%; Score 180.5; DB 1; Length 473;
  Best Local Similarity 23.8%; Pred. No. 9.4e-05;
  Matches 65; Conservative 41; Mismatches 84; Indels 83; Gaps 7;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|:::| :   :     | :||::||:     | ||::|: ||:|  :|  |: ||:|
Db          5 ALELKNEANKFLKEGHIVQAIDLYTKAIELDSTNAILYSNRSLAHLKSEDYGLAINDASK 64

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||| ||   |||:|:  | | : : : |        : |  |    : |:||:: :
Db         65 AIECDPEYAKAYFRRATAHIAIFQPKEAVGDFRKALALAPSDPAARKKLRECEQLVKRIR 124

Qy        123 SQAPVKNVE---ATVAATIED----------------------KEDFTN--------- 146
                | : | |          |||                      || |
Db        125 FQEAIHNTEPPSPLANINIEDMDIPSDYDGVILEKQITKEFVEDMKERFCQGKKLPLKFA 184

Qy        147 ----------MENTPPVIE------------------------------PPS---KPKYR 163
                      :| || :|:                              |||   |  :
Db        185 YSILRDLKELLEKTPSLIDIPVKGDETLVICGDTHGQYFDLLNIFKLHGPPSPTNKYLFN 244

Qy        164 HDYYNS---ATEVVLTIFA-KGVPADSVVIDFG 192
              |: :    :|||  |::| | :  |:| |: |
Db        245 GDFVDRGSWSTEVAFTLYAYKLLYPDAVFINRG 277
```

RESULT 14
A32567
stress-induced protein STI1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: heat shock protein STI1; protein O2657; protein YOR027w
C;Species: Saccharomyces cerevisiae
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: A32567; S54633; S66893
R;Nicolet, C.M.; Craig, E.A.
Mol. Cell. Biol. 9, 3638-3646, 1989
A;Title: Isolation and characterization of STI1, a stress-inducible gene from Saccharomy
A;Reference number: A32567; MUID:89384590; PMID:2674681
A;Accession: A32567
A;Molecule type: DNA
A;Residues: 1-589 <NIC>
A;Cross-references: UNIPROT:P15705; GB:M28486; NID:g172765; PIDN:AAA35121.1; PID:g172766
R;de Haan, M.; Maarse, A.C.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54617
A;Accession: S54633
A;Molecule type: DNA
A;Residues: 1-589 <DEH>
A;Cross-references: EMBL:X87331; NID:g1041652; PID:g829137
R;de Haan, M.; Grivell, L.A.; Maarse, A.C.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66877
A;Accession: S66893
A;Molecule type: DNA
A;Residues: 1-589 <DEW>
A;Cross-references: EMBL:Z74935; NID:g1420138; PID:e252326; PID:g1420139; MIPS:YOR027w
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:STI1
A;Cross-references: SGD:S0005553; MIPS:YOR027w
A;Map position: 15R
C;Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat ho
F;5-38/Domain: tetratricopeptide repeat homology <TT1>
F;40-73/Domain: tetratricopeptide repeat homology <TT2>
F;74-107/Domain: tetratricopeptide repeat homology <TT3>

  Query Match 9.5%; Score 177.5; DB 2; Length 589;
  Best Local Similarity 33.3%; Pred. No. 0.0002;
  Matches 52; Conservative 22; Mismatches 59; Indels 23; Gaps 4;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :   : || |    |:  | : |:: |    |  |  |::|| |  || :: ||:|| ||
Db        396 AEEARLEGKEYFTKSDWPNAVKAYTEMIKRAPEDARGYSNRAAALAKLMSFPEAIADCNK 455

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALEL----------GSS--------YASGD 104
              ||| ||   :|| ||  | | ::|| :|   |:           |||         | :
Db        456 AIEKDPNFVRAYIRKATAQIAVKEYASALETLDAARTKDAEVNNGSSAREIDQLYYKASQ 515

Qy        105 SRFARLLKECDERIAEESSQAPVKNVEATVAATIED 140
              ||          ||: |  :|: |  ||| ::|
Db        516 QRFQ---PGTSNETPEETYQRAMKDPE--VAAIMQD 546
```

RESULT 15
S56658
stress-induced protein sti1 - soybean
C;Species: Glycine max (soybean)
C;Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 16-Aug-2004
C;Accession: S56658; S56659; S71763
R;Torres, J.H.; Chatellard, P.; Stutz, E.
Plant Mol. Biol. 27, 1221-1226, 1995
A;Title: Isolation and characterization of gmsti, a stress-inducible gene from soybean (
A;Reference number: S56658; MUID:95284372; PMID:7766904
A;Accession: S56658
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-569 <TOR>
A;Cross-references: UNIPROT:Q43468; EMBL:X79770
A;Accession: S56659
A;Molecule type: DNA
A;Residues: 417-569 <TOW>
A;Cross-references: EMBL:X79770
R;Stutz, E.
submitted to the EMBL Data Library, June 1994
A;Reference number: S71763
A;Accession: S71763
A;Molecule type: mRNA
A;Residues: 1-24,'F',26-569 <STU>
A;Cross-references: EMBL:X79770; NID:g872115; PID:g872116
C;Genetics:
A;Gene: sti1
A;Introns: 491/2; 513/3; 533/3
A;Note: the list of introns is incomplete

C;Superfamily: tetratricopeptide repeat homology
F;321-354/Domain: tetratricopeptide repeat homology <TT1>
F;381-414/Domain: tetratricopeptide repeat homology <TT2>
F;415-448/Domain: tetratricopeptide repeat homology <TT3>
F;449-482/Domain: tetratricopeptide repeat homology <TT4>

Query Match 9.5%; Score 176; DB 2; Length 569;
Best Local Similarity 33.3%; Pred. No. 0.00024;
Matches 40; Conservative 17; Mismatches 63; Indels 0; Gaps 0;

```
Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :   |  | |    :  ||: |::||   |  |  |::||  : |||   | : || |
Db        381 ADEAREKGNELFKQQKYPEATKHYTEAIKRNPKDAKAYSNRAACYTKLGAMPEGLKDAEK 440

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||   |. | ||||    ::||  |      |  :   :      :: | |:| : |
Db        441 CIELDPTFSKGYTRKGAVQFSMKEYDKALETYREGLKHDPNNQELLDGIRRCVEQINKAS 500
```

Search completed: April 20, 2005, 18:35:07
Job time : 43 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:17:45 ; Search time 178 Seconds
(without alignments)
1038.543 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1594.5	85.7	373	2	Q8W516	Q8w516 hordeum vul
2	1592.5	85.6	367	2	Q9FPB1	Q9fpb1 oryza sativ
3	1589.5	85.4	367	2	Q9SE32	Q9se32 oryza sativ
4	1243	66.8	370	2	Q84UV7	Q84uv7 nicotiana b
5	1142	61.4	350	2	Q8W515	Q8w515 arabidopsis
6	1138	61.1	354	2	Q704T1	Q704t1 brassica ol
7	1135	61.0	358	2	Q9SUT5	Q9sut5 arabidopsis
8	1134	60.9	350	2	Q9SUR9	Q9sur9 arabidopsis
9	1133.5	60.9	327	2	Q9SM53	Q9sm53 rubus idaeu
10	1125.5	60.5	355	2	Q704T2	Q704t2 brassica ol
11	870	46.7	273	2	Q84LL4	Q84ll4 arabidopsis
12	615	33.0	333	1	SUGT_HUMAN	Q9y2z0 homo sapien
13	608.5	32.7	331	2	Q6IP01	Q6ip01 xenopus lae
14	602.5	32.4	211	2	Q8LKF8	Q8lkf8 nicotiana b
15	602.5	32.4	211	2	Q8LKF9	Q8lkf9 nicotiana b
16	599.5	32.2	336	1	SUGT_MOUSE	Q9cx34 mus musculu
17	594	31.9	365	2	Q6VXY6	Q6vxy6 homo sapien
18	423	22.7	375	2	Q6BXK8	Q6bxk8 debaryomyce
19	383	20.6	198	2	Q18949	Q18949 caenorhabdi
20	370.5	19.9	377	2	Q6FKN1	Q6fkn1 candida gla
21	364.5	19.6	395	1	SGT1_YEAST	Q08446 saccharomyc
22	364.5	19.6	395	2	Q6Q5I4	Q6q5i4 saccharomyc
23	348.5	18.7	379	2	Q6CJS5	Q6cjs5 kluyveromyc
24	341	18.3	379	1	GIT7_SCHPO	O59709 schizosacch
25	334.5	18.0	466	2	Q9P6Y5	Q9p6y5 neurospora
26	329	17.7	221	2	Q9N8L0	Q9n8l0 trypanosoma
27	321.5	17.3	204	2	Q7Q5W9	Q7q5w9 anopheles g
28	317	17.0	178	2	Q9VHT3	Q9vht3 drosophila
29	314	16.9	178	2	Q8SY87	Q8sy87 drosophila
30	313.5	16.8	247	2	Q7TQ12	Q7tq12 rattus norv
31	306.5	16.5	462	2	Q7RV15	Q7rv15 neurospora
32	300	16.1	368	2	Q75EE3	Q75ee3 ashbya goss
33	252.5	13.6	519	2	Q6C1J0	Q6c1j0 yarrowia li
34	224	12.0	485	2	Q8H1H4	Q8h1h4 lycopersico
35	221.5	11.9	484	2	Q9SJH5	Q9sjh5 arabidopsis
36	219.5	11.8	538	2	Q84XU2	Q84xu2 arabidopsis
37	209.5	11.3	556	2	Q84K11	Q84k11 lycopersico
38	208.5	11.2	479	2	Q6CFH3	Q6cfh3 yarrowia li
39	200.5	10.8	328	2	Q8RXN1	Q8rxn1 arabidopsis
40	200	10.7	232	2	Q6Q0C2	Q6q0c2 nicotiana b
41	198	10.6	343	2	Q6FVD6	Q6fvd6 candida gla
42	197.5	10.6	571	2	Q6BPQ7	Q6bpq7 debaryomyce
43	194.5	10.5	564	2	P90647	P90647 acanthamoeb
44	194	10.4	492	2	O42205	O42205 xenopus lae
45	194	10.4	493	2	Q6GPS6	Q6gps6 xenopus lae

ALIGNMENTS

RESULT 1
Q8W516
ID Q8W516 PRELIMINARY; PRT; 373 AA.
AC Q8W516;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SGT1.
GN Name=SGT1;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;
RA Azevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shirasu K.,
RA Schulze-Lefert P.;
RT "The RAR1 interactor SGT1, an essential component of R gene-triggered
RT disease resistance.";
RL Science 295:2073-2076(2002).
DR EMBL; AF439974; AAL33610.1; -.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF00515; TPR_1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 373 AA; 40997 MW; 74A6E2C77D0F9EAD CRC64;

Query Match 85.7%; Score 1594.5; DB 2; Length 373;
Best Local Similarity 83.0%; Pred. No. 1e-95;
Matches 307; Conservative 26; Mismatches 26; Indels 11; Gaps 2;

```
Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||||||||||||| |||:|||:||||||:|||||||||||||:|||||||||
Db          5 AASDLESKAKEAFVDDDFELAAELYTQAIEAGPATAELYADRAQAHIKLGSYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||:|||||||||||||||:|||||||||| ||:||||:|||||
Db         65 KAIELDPSMHKAYLRKGSACIKLEEYQTAKAALEVGSSSYASGDSRFTRLMKECDDRIAEE 124

Qy        122 SSQAPVKNVEATVA----------ATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSAT 171
              :|||||| | ||          | : :|  ||||  | :| ||||||||||: |
Db        125 ASQAPVKNAAAAVAPATSSGATTVVTEAEDQDGENMENAQPTVEVPSKPKYRHDYYNTPT 184

Qy        172 EVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTK 231
              ||||||||||||||||:||||||||||:||||||||||||||||:|:|||| |||||
Db        185 EVVLTIFAKGVPADSVVVDFGEQMLSVSIELPGEEPYHFQPRLFSKIVPDKCKYTVLSTK 244

Qy        232 VEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKE 291
              ||||||||| |||||:|||:|:||  ||||: |||:| |||||||:||||||||||||:
```

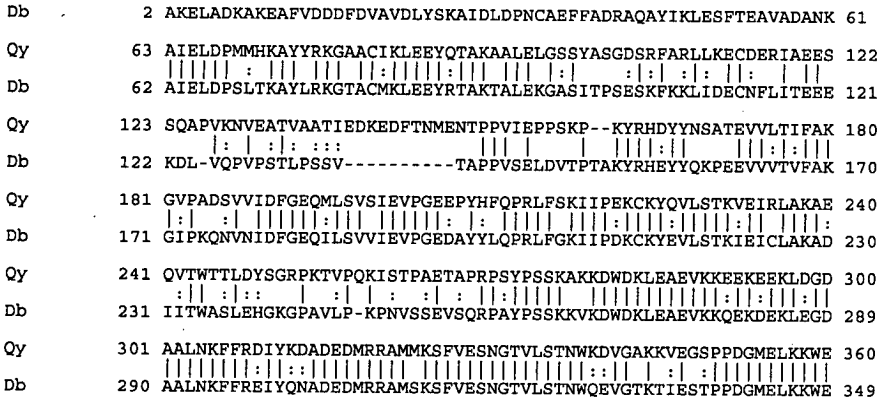
```
Db        245 VEIRLAKAEPVTWTSLDYTGKPK-APQKINVPAESAQRPSYPSSKSKKDWDKLEAEVKKQ 303

Qy        292 EKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPP 351
              ||:|||||||||||||:|| |||||||||||||||||||||||||||||| | ||||||
Db        304 EKDEKLDGDAALNKFFREIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGKKTVEGSPP 363

Qy        352 DGMELKKWEY 361
              ||||||||||
Db        364 DGMELKKWEY 373


RESULT 2
Q9FPB1
ID   Q9FPB1      PRELIMINARY;      PRT;   367 AA.
AC   Q9FPB1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Sgt1.
GN   Name=P0006C01.2;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12447438; DOI=10.1038/nature01184;
RA   Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA   Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA   Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA   Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA   Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA   Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA   Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA   Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA   Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA   Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA   Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA   Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA   Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA   Yano M., Jiang J., Gojobori T.;
RT   "The genome sequence and structure of rice chromosome 1.";
RL   Nature 420:312-316(2002).
DR   EMBL; AP002744; BAB19060.1; -.
DR   Gramene; Q9FPB1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 1.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   367 AA;  40938 MW;  885E7758A754FED1 CRC64;

  Query Match             85.6%;  Score 1592.5;  DB 2;  Length 367;
  Best Local Similarity   84.1%;  Pred. No. 1.4e-95;
  Matches  307;  Conservative   27;  Mismatches   24;  Indels    7;  Gaps    2;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||| |||||||||| |||:|||:| ||||:|||||||||||||||||||||||
Db          5 AASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYADRAQAHIKLGNYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||||||:|||||||||||| |:|||||| ||:|||||||||||
Db         65 KAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYSFASGDSRFTRLMKECDERIAEE 124

Qy        122 SSQAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLT 176
```

```
              |: ||| |      :||: :|:|:|  ||:||||::|   |||||||:||||||||||
Db        125 LSEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              ||||||||::||:||||||||||||||||||||||||||||||| :||||||||||||
Db        183 IFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKSRYQVLSTKVEIRL 242

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEK 296
              |||||:|||:|||  :|| |||||  |||:| ||||||||:||||||||||||||||||
Db        243 AKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRPSYPSSKSKKDWDKLEAEVKKEEKEEK 302

Qy        297 LDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEL 356
              |:|||||||||||| ||||||||||||||||||||||||||:||||||||||||||
Db        303 LEGDAALNKFFRDIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVEGSPPDGMEL 362

Qy        357 KKWEY 361
              |||||
Db        363 KKWEY 367


RESULT 3
Q9SE32
ID   Q9SE32      PRELIMINARY;      PRT;   367 AA.
AC   Q9SE32;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Sgt1.
GN   Name=Sgt1;
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20035956; PubMed=10571178; DOI=10.1016/S0092-8674(00)81522-6;
RA   Shirasu K., Lahaye T., Tan M.W., Zhou F., Azevedo C.,
RA   Schulze-Lefert P.;
RT   "A novel class of eukaryotic zinc-binding proteins is required for
RT   disease resistance signaling in barley and development in C.
RT   elegans.";
RL   Cell 99:355-366(1999).
DR   EMBL; AF192467; AAF18438.1; -.
DR   Gramene; Q9SE32; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 1.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   367 AA;  40952 MW;  885F320DF203FED1 CRC64;

  Query Match             85.4%;  Score 1589.5;  DB 2;  Length 367;
  Best Local Similarity   83.8%;  Pred. No. 2.1e-95;
  Matches  306;  Conservative   28;  Mismatches   24;  Indels    7;  Gaps    2;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||| |||||||||| |||:|||:| ||||:|||||||||||||||||||||||
Db          5 AASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYADRAQAHIKLGNYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||||||:|||||||||||| |:|||||| ||:|||||||||||
Db         65 KAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYSFASGDSRFTRLMKECDERIAEE 124

Qy        122 SSQAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLT 176
              :: |||  |      :||: :|:|:|  ||:||||::|   |||||||:||||||||||
```

```
Db        125 LTEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              ||||||||::||:||||||||||||||||||||||||||||||| :|||||||||||||
Db        183 IFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKSRYQVLSTKVEIRL 242

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEK 296
              |||||:|||:|||  :|| |||||  |||:| ||||||||:||||||||||||||||||
Db        243 AKAEQITWTSLDYDKKPKAVPQKIIPPAESAQRPSYPSSKSKKDWDKLEAEVKKEEKEEK 302

Qy        297 LDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEL 356
              |:|||||||||||||| ||||||||||||||||||||||||||:|||||||||||||||
Db        303 LEGDAALNKFFRDIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVEGSPPDGMEL 362

Qy        357 KKWEY 361
              |||||
Db        363 KKWEY 367


RESULT 4
Q84UV7
ID   Q84UV7      PRELIMINARY;      PRT;   370 AA.
AC   Q84UV7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SGT1.
OS   Nicotiana benthamiana.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4100;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22115173; PubMed=12119369;
RA   Liu Y., Schiff M., Serino G., Deng X.W., Dinesh-Kumar S.P.;
RT   "Role of SCF ubiquitin-ligase and the COP9 signalosome in the N gene-
RT   mediated resistance response to Tobacco mosaic virus.";
RL   Plant Cell 14:1483-1496(2002).
DR   EMBL; AF494083; AAO85509.1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF00515; TPR_1; 1.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   370 AA;  41240 MW;  88EF4C3408F5121F CRC64;

  Query Match             66.8%;  Score 1243;  DB 2;  Length 370;
  Best Local Similarity   66.2%;  Pred. No. 7.5e-73;
  Matches  245;  Conservative   48;  Mismatches   63;  Indels   14;  Gaps    5;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |||||::|||||:|| |||| :||:|||    |  |:|:||||||:|||  :|||| ||||
Db          2 ASDLETRAKEAFIDDHFELAVDLYTQAIAMTPKNAELFADRAQANIKLNYFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||| | ||| ||| ||:|||||||||||||| |:| |  :||| :|:||||||||:
Db         62 AIELDPSMSKAYLRKGLACMKLEEYQTAKAALETGASLAPAESRFTKLIKECDERIAEEA 121

Qy        123 SQAPVKNVEATVAATIED--KEDFTNMENTPPVIEP---------PSKPKYRHDYYNSAT 171
                : | ::|: |    :     |   |:   |   :|          ::|||||::|
Db        122 GELPNQSVDKTSGNVVAPPASESLDNVAVAPKDAQPTVNLSYQGSAARPKYRHEFYQKPE 181

Qy        172 EVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTK 231
              |||:||||||:|| :||:|||||:|||||:|||:| | |||||| || | ||:|:|:|||
Db        182 EVVVTIFAKGIPAKNVVVDFGEQILSVSIDVPGDETYSFQPRLFGKITPAKCRYEVMSTK 241

Qy        232 VEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKK-DWDKLEAEVKK 290
              :|||||||| : ||:|:|:| |   | |: :  :: |||||||||| :  ||||||||||
Db        242 IEIRLAKAEPLHWTSLEYT-RASAVVQRPNVSSD-APRPSYPSSKLRHVDWDKLEAEVKK 299

Qy        291 EEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSP 350
              |||:|||||||||||||||||||| ||||||||||||||||||:|| |||||||
Db        300 EEKDEKLDGDAALNKFFRDIYKDADEDTRRAMMKSFVESNGTVLSTNWKEVGTKKVEGSP 359

Qy        351 PDGMELKKWE 360
              ||||||||||
Db        360 PDGMELKKWE 369


RESULT 5
Q8W515
ID   Q8W515      PRELIMINARY;      PRT;   350 AA.
AC   Q8W515;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SGT1a.
GN   Name=SGT1a;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;
RA   Azevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shirasu K.,
RA   Schulze-Lefert P.;
RT   "The RAR1 interactor SGT1, an essential component of R gene-triggered
RT   disease resistance.";
RL   Science 295:2073-2076(2002).
DR   EMBL; AF439975; AAL33611.1; -.
DR   HSSP; P31948; 1ELW.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 2.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 2.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   350 AA;  39277 MW;  D23DBE990C817FB4 CRC64;

  Query Match             61.4%;  Score 1142;  DB 2;  Length 350;
  Best Local Similarity   61.7%;  Pred. No. 2.6e-66;
  Matches  222;  Conservative   56;  Mismatches   68;  Indels   14;  Gaps    4;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  |||||||||||::| :|||:|||  |  |: :|||||:|||  ::||||||||||
Db          2 AKELADKAKEAFVDDDFDVAVDLYSKAIDLDPNCAEFFADRAQAYIKLESFTEAVADANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              |||||| : ||| ||| ||:|||||:||| ||| |:|    :|:| :|: ||:  | ||
Db         62 AIELDPSLTKAYLRKGTACMKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEE 121

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAK 180
                |: | :|: :::          ||| |    | ||||:||    |||:|:|||
Db        122 KDL-VQPVPSTLPSSV----------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAK 170

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              |:|  :| ||||||:||| ||||||: |: ||||| ||||:||||:|||||:|||||||:
Db        171 GIPKQNVNIDFGEQILSVVIEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEIRLAKAD 230

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
              :|| :|::   |  :| | :  :| : ||:|||||  |||||||||||:||:|||:||
Db        231 IITWASLEHGKGPAVLP-KPNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGD 289

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||:||::||||||||| ||||||||||||||::|| | :| :|||||||||||
Db        290 AALNKFFREIYQNADEDMRRAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 349
```

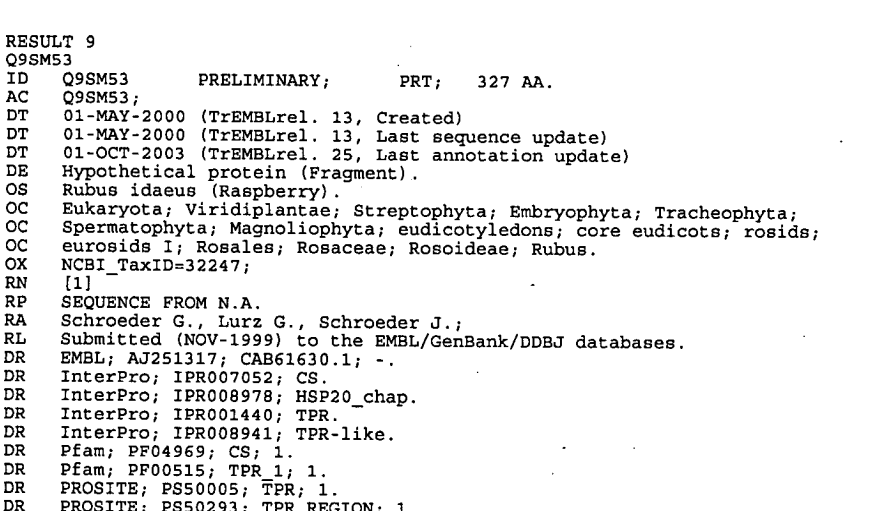
```
RESULT 6
Q704T1
ID   Q704T1      PRELIMINARY;      PRT;   354 AA.
AC   Q704T1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   SGT1-like protein.
GN   Name=sgt1b-2;
OS   Brassica oleracea (Cauliflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3712;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Yemm A.A., Holub E., Tor M.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ620883; CAF06581.1; -.
DR   HSSP; P53041; 1A17.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 2.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   354 AA;  39041 MW;  B97D6F8A92285A97 CRC64;

  Query Match             61.1%;  Score 1138;  DB 2;  Length 354;
  Best Local Similarity   62.4%;  Pred. No. 4.8e-66;
  Matches  224;  Conservative   53;  Mismatches   76;  Indels    6;  Gaps    4;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              ||:|  |||||:||||::| :|||:|||  |: |  :||||||:||| |:|||||||||
Db          2 ASELAEKAKEAFLDDDFDVAADLYSKAIDLDPSCASFFADRAQANIKLLNFTEAVADANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||||||:||||| ||||||| |:| |  :|:| ::: ||:  ||||
Db         62 AIELEPTLAKAYLRKGAACMKLEEYATAKAALEKGASVAPNESKFEKMIDECNLLIAEEE 121

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182
                |: |  |: ::          : ||    |:|| :||::|    || : |||||:
Db        122 KDL-VQQVPPTLPSS--STTPLAIAADAPPA--APAKPMFRHEFYQKPEEVGVAIFAKGI 176

Qy        183 PADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQV 242
              |  :| ::||:|:||| |:| ||| |||||||| ||||||||:|:||||||||||||| |
Db        177 PKQNVNVEFGDQILSVVIDVAGEEAYHFQPRLFGKIIPEKCRYEVLSTKVEIRLAKAEIV 236

Qy        243 TWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAA 302
              || :|:| |: : :  | :  :  : || |||||  |||||||||||:||:||||||||
Db        237 TWASLEY-GKGQALLPKPNVASAVSQRPVYPSSKPGKDWDKLEAEVKKQEKDEKLDGDAA 295

Qy        303 LNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWEY 361
              :|||| |||: |||||||||| ||| ||||||||||||:|| |||| :||||||||||
Db        296 MNKFFSDIYQSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGMELKKWEY 354


RESULT 7
Q9SUT5
ID   Q9SUT5      PRELIMINARY;      PRT;   358 AA.
AC   Q9SUT5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein AT4g11260 (SGT1a) (AT4g11260/F8L21_50) (SGT1b)
DE   (SGT1-like protein).
GN   Name=AT4g11260; Synonyms=At4g11260, SGT1b, sgt1a;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA   Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22088475; PubMed=12093376;
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21893744; PubMed=11847307; DOI=10.1126/science.1067554;
RA   Azevedo C., Sadanandom A., Kitagawa K., Freialdenhoven A., Shirasu K.,
RA   Schulze-Lefert P.;
RT   "The RAR1 interactor SGT1, an essential component of R gene-triggered
RT   disease resistance.";
RL   Science 295:2073-2076(2002).
RN   [9]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22882917; PubMed=14504394; DOI=10.1073/pnas.2035132100;
RA   Pawlowski J., Holzmann M., Berney C., Fahrni J., Gooday A.J.,
RA   Cedhagen T., Habura A., Bowser S.S.;
RT   "The evolution of early Foraminifera.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:11494-11498(2003).
RN   [10]
RP   SEQUENCE FROM N.A.
RA   Shen W.H.;
```

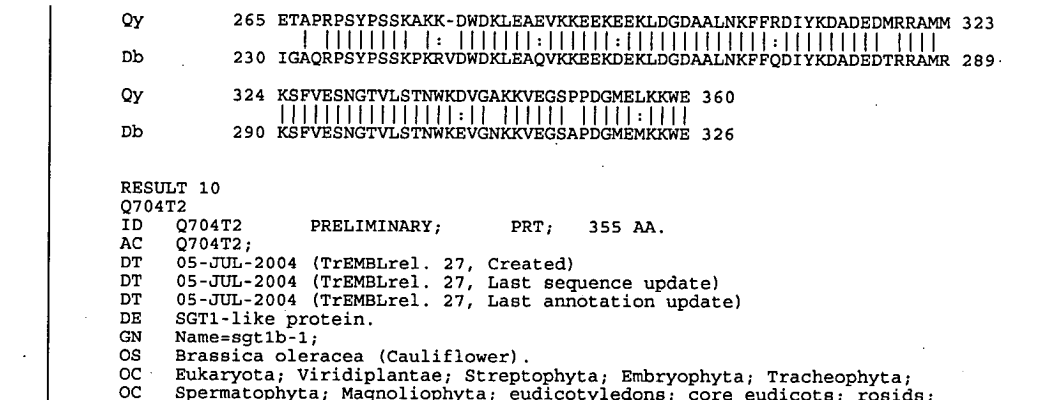
```
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA   Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA   Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA   Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA   Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL096882; CAB51410.1; -.
DR   EMBL; AL161531; CAB81227.1; -.
DR   EMBL; AY085464; AAM62690.1; -.
DR   EMBL; AF428340; AAL16270.1; -.
DR   EMBL; AJ318019; CAC85266.1; -.
DR   EMBL; AY150487; AAN12904.1; -.
DR   EMBL; AF439976; AAL33612.1; -.
DR   EMBL; AF370229; AAK44044.1; -.
DR   PIR; T13017; T13017.
DR   HSSP; P31948; 1ELW.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 2.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 2.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Hypothetical protein; Repeat; TPR repeat.
SQ   SEQUENCE   358 AA;  39762 MW;  711172817DD349C3 CRC64;

  Query Match           61.0%;  Score 1135;  DB 2;  Length 358;
  Best Local Similarity  61.0%;  Pred. No. 7.6e-66;
  Matches  224;  Conservative   49;  Mismatches   76;  Indels   18;  Gaps    4;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||:||||::| :|||:|||  |  |  :||||||:||: |:|||| ||||
Db          2 AKELAEKAKEAFLDDDFDVAVDLYSKAIDLDPNCAAFFADRAQANIKIDNFTEAVVDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:| : ||| ||| ||:||||| ||||||| |:| |  : :| ::: ||| |||||
Db         62 AIELEPTLAKAYLRKGTACMKLEEYSTAKAALEKGASVAPNEPKFKKMIDECDLRIAEEE 121

Qy        123 SQ------APVKNVEATVAATIEDKEDFTNMENTPPVIEP--PSKPKYRHDYYNSATEVV 174
                    : :   |  ||  |           |||  |  |:|| :||::|    | |
Db        122 KDLVQPMPPSLPSSSTTPLATEAD---------APPVPIPAAPAKPMFRHEFYQKPEEAV 172

Qy        175 LTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEI 234
              :||||| || ::| ::|||:||| |:| ||| || ||||| |||||||:::||||||||
Db        173 VTIFAKKVPKENVTVEFGEQILSVVIDVAGEEAYHLQPRLFGKIIPEKCRFEVLSTKVEI 232

Qy        235 RLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKE 294
              |||||| :|| :|:| |: ::|  | :  :  : || |||||  ||||||||||||:||:
Db        233 RLAKAEIITWASLEY-GKGQSVLPKPNVSSALSQRPVYPSSKPAKDWDKLEAEVKKQEKD 291

Qy        295 EKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGM 354
              ||||||||:|||| |||  |||||||||| ||| ||||||||||||:|| |||| :|||||
Db        292 EKLDGDAAMNKFFSDIYSSADEDMRRAMNKSFAESNGTVLSTNWKEVGTKKVESTPPDGM 351

Qy        355 ELKKWEY 361
              |||||||
Db        352 ELKKWEY 358


RESULT 8
Q9SUR9
ID   Q9SUR9      PRELIMINARY;      PRT;   350 AA.
AC   Q9SUR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Phosphatase like protein.
GN   Name=F9D16.40; Synonyms=AT4g23570, At4g23570;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Wedler H., Wedler E., Wambutt R., Hoheisel J., Mewes H.W.,
RA   Mayer K.F.X., Schueller C.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Wedler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035394; CAA23023.1; -.
DR   EMBL; AY063978; AAL36334.1; -.
DR   EMBL; AY096395; AAM20035.1; -.
DR   EMBL; AL161559; CAB79312.1; -.
DR   PIR; T05589; T05589.
DR   HSSP; P31948; 1ELW.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 2.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50005; TPR; 2.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   350 AA;  39224 MW;  D23DBE83BC8ACFA5 CRC64;

  Query Match           60.9%;  Score 1134;  DB 2;  Length 350;
  Best Local Similarity  61.4%;  Pred. No. 8.6e-66;
  Matches  221;  Conservative   56;  Mismatches   69;  Indels   14;  Gaps    4;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              | :|  ||||||||||||::| :|||:|||  |  |: :||||||:||| ::|||||||||
```



RESULT 9
Q9SM53
ID Q9SM53 PRELIMINARY; PRT; 327 AA.
AC Q9SM53;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rubus idaeus (Raspberry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Rosoideae; Rubus.
OX NCBI_TaxID=32247;
RN [1]
RP SEQUENCE FROM N.A.
RA Schroeder G., Lurz G., Schroeder J.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ251317; CAB61630.1; -.
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF00515; TPR_1; 1.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 327 AA; 36617 MW; A7CB03A42D2032B8 CRC64;

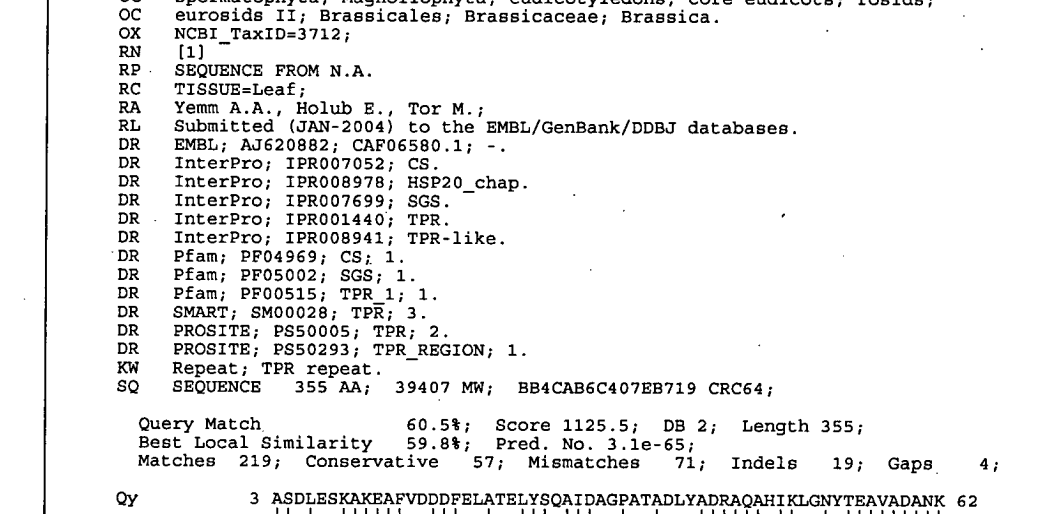
Query Match 60.9%; Score 1133.5; DB 2; Length 327;
Best Local Similarity 66.2%; Pred. No. 8.5e-66;
Matches 223; Conservative 44; Mismatches 49; Indels 21; Gaps 5;





RESULT 10
Q704T2
ID Q704T2 PRELIMINARY; PRT; 355 AA.
AC Q704T2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SGT1-like protein.
GN Name=sgt1b-1;
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Yemm A.A., Holub E., Tor M.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ620882; CAF06580.1; -.
DR InterPro; IPR007052; CS.
DR InterPro; IPR008978; HSP20_chap.
DR InterPro; IPR007699; SGS.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF04969; CS; 1.
DR Pfam; PF05002; SGS; 1.
DR Pfam; PF00515; TPR_1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50005; TPR; 2.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.
SQ SEQUENCE 355 AA; 39407 MW; BB4CAB6C407EB719 CRC64;

Query Match 60.5%; Score 1125.5; DB 2; Length 355;
Best Local Similarity 59.8%; Pred. No. 3.1e-65;
Matches 219; Conservative 57; Mismatches 71; Indels 19; Gaps 4;



```
Db        350 LKKWEY 355


RESULT 11
Q84LL4
ID   Q84LL4      PRELIMINARY;      PRT;   273 AA.
AC   Q84LL4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SGT1-like protein (Fragment).
GN   Name=sgt1b;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22882917; PubMed=14504394; DOI=10.1073/pnas.2035132100;
RA   Pawlowski J., Holzmann M., Berney C., Fahrni J., Gooday A.J.,
RA   Cedhagen T., Habura A., Bowser S.S.;
RT   "The evolution of early Foraminifera.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:11494-11498(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shen W.H.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ318020; CAC85267.1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   273 AA;  30721 MW;  0D8A25034F64DC59 CRC64;

  Query Match             46.7%;  Score 870;  DB 2;  Length 273;
  Best Local Similarity   59.7%;  Pred. No. 9.3e-49;
  Matches  169;  Conservative   45;  Mismatches   55;  Indels   14;  Gaps    4;

Qy         80 ACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIE 139
              ||:|||||:||| ||| |:|    :|:| :|: ||:  | ||      |: | :|: :::
Db          2 ACMKLEEYRTAKTALEKGASITPSESKFKKLIDECNFLITEEEKDL-VQPVPSTLPSSV- 59

Qy        140 DKEDFTNMENTPPVIEPPSKP--KYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLS 197
                     ||| |    |  ||||:||    |||:|:||||:|  :| |||||:||
Db         60 ---------TAPPVSELDVTPTAKYRHEYYQKPEEVVVTVFAKGIPKQNVNIDFGEQILS 110

Qy        198 VSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVP 257
              | ||||||: |: ||||| ||||:||||:|||||:|| ||||: :|| :|::   |  :|
Db        111 VVIEVPGEDAYYLQPRLFGKIIPDKCKYEVLSTKIEICLAKADIITWASLEHGKGPAVLP 170

Qy        258 QKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADED 317
               | :  :| : ||:|||||  |||||||||||||:||:|||:||||||||||:||::||||
Db        171 -KPNVSSEVSQRPAYPSSKKVKDWDKLEAEVKKQEKDEKLEGDAALNKFFREIYQNADED 229

Qy        318 MRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||| |||||||||||||::|| | :| :||||||||||
Db        230 MRRAMSKSFVESNGTVLSTNWQEVGTKTIESTPPDGMELKKWE 272


RESULT 12
SUGT_HUMAN
ID   SUGT_HUMAN      STANDARD;      PRT;   333 AA.
AC   Q9Y2Z0;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Suppressor of G2 allele of SKP1 homolog (Sgt1) (Putative 40-6-3
DE   protein).
GN   Name=SUGT1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SUBUNIT.
RX   MEDLINE=99374050; PubMed=10445024; DOI=10.1016/S1097-2765(00)80184-7;
RA   Kitagawa K., Skowyra D., Elledge S.J., Harper J.W., Hieter P.;
RT   "SGT1 encodes an essential component of the yeast kinetochore assembly
RT   pathway and a novel subunit of the SCF ubiquitin ligase complex.";
RL   Mol. Cell 4:21-33(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Schmidt T.;
RL   Thesis (2001), University of Goettingen, Germany.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RT   "Cloning of human full-length CDSs in BD Creator(TM) system donor
RT   vector.";
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   INTERACTION WITH S100A6.
RX   MEDLINE=22741750; PubMed=12746458; DOI=10.1074/jbc.M211518200;
RA   Nowotny M., Spiechowicz M., Jastrzebska B., Filipek A., Kitagawa K.,
RA   Kuznicki J.;
RT   "Calcium-regulated interaction of Sgt1 with S100A6 (calcyclin) and
RT   other S100 proteins.";
RL   J. Biol. Chem. 278:26923-26928(2003).
CC   -!- FUNCTION: May play a role in ubiquitination and subsequent
CC       proteosomal degradation of target proteins.
CC   -!- SUBUNIT: Probably associates with SCF (SKP1-CUL1-F-box protein)
CC       complex through interaction with SKP1. Interacts with S100A6.
CC   -!- SIMILARITY: Contains 1 CS domain.
CC   -!- SIMILARITY: Contains 1 SGS domain.
CC   -!- SIMILARITY: Contains 3 TPR repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF132856; AAD30062.1; -.
DR   EMBL; AJ344097; CAC51433.1; -.
DR   EMBL; BT009798; AAP88800.1; -.
```

```
DR   EMBL; BC000911; AAH00911.1; -.
DR   Genew; HGNC:16987; SUGT1.
DR   MIM; 604098; -.
DR   GO; GO:0000776; C:kinetochore; TAS.
DR   GO; GO:0000151; C:ubiquitin ligase complex; TAS.
DR   GO; GO:0000069; P:centromere/kinetochore complex maturation; TAS.
DR   GO; GO:0007067; P:mitosis; TAS.
DR   GO; GO:0000074; P:regulation of cell cycle; TAS.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR; 3.
DR   PROSITE; PS50005; TPR; 3.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat; Ubl conjugation pathway.
FT   REPEAT       11     44       TPR 1.
FT   REPEAT       45     78       TPR 2.
FT   REPEAT       79    112       TPR 3.
FT   DOMAIN      138    225       CS.
FT   DOMAIN      252    333       SGS.
SQ   SEQUENCE   333 AA;  37804 MW;  14DE069A1623DB53 CRC64;

  Query Match           33.0%;  Score 615;  DB 1;  Length 333;
  Best Local Similarity   38.7%;  Pred. No. 4.6e-32;
  Matches  135;  Conservative   62;  Mismatches  116;  Indels   36;  Gaps    5;

Qy         12 EAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMH 71
              :| :|:| : | |  ::|::  |  |  |  ||  || ||||  ||||| |::||:|
Db         20 DALIDEDPQAALEELTKALEQKPDDAQYYCQRAYCHILLGNYCVAVADAKKSLELNPNNS 79

Qy         72 KAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVE 131
              |  |||      : |  |      |    | |: |:  :| | |  |:  |:: |
Db         80 TAMLRKGICEYHEKNYAAALETFTEGQKLDSADANFSVWIKRCQE--AQNGSESEV---- 133

Qy        132 ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDF 191
                        :|:            : | ::|:| : ::||:|:  | |  : | ::|
Db        134 ------------WTH------------QSKIKYDWYQTESQVVITLMIKNVQKNDVNVEF 169

Qy        192 GEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSG 251
              |: ||  :::|  | |: :  |   ||||:  ::||||:||:| | | | |  |:  |
Db        170 SEKELSALVKLPSGEDYNLKLELLHPIIPEQSTFKVLSTKIEIKLKKPEAVRWEKLEGQG 229

Qy        252 RPKTVPQKISTPAETAPRPSYPSSKA-KKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDI 310
              |  | ::       :  ||||    ::|||| |:|:||| |||:||||||: |: |
Db        230 DVPTPKQFVAD-----VKNLYPSSSPYTRNWDKLVGEIKEEEKNEKLEGDAALNRLFQQI 284

Qy        311 YKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 359
              | |  ::::||| |||:|| |||||||| ||| :||| :||| || ||:
Db        285 YSDGSDEVKRAMNKSFMESGGTVLSTNWSDVGKRKVEINPPDDMEWKKY 333


RESULT 13
Q6IP01
ID   Q6IP01      PRELIMINARY;      PRT;   331 AA.
AC   Q6IP01;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC79143 protein.
GN   Name=MGC79143;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072118; AAH72118.1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR007699; SGS.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF05002; SGS; 1.
DR   Pfam; PF00515; TPR_1; 2.
DR   SMART; SM00028; TPR; 2.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Repeat; TPR repeat.
SQ   SEQUENCE   331 AA;  37924 MW;  4C910A75B33D81AA CRC64;

  Query Match           32.7%;  Score 608.5;  DB 2;  Length 331;
  Best Local Similarity   37.9%;  Pred. No. 1.2e-31;
  Matches  132;  Conservative   63;  Mismatches  122;  Indels   31;  Gaps    5;

Qy         12 EAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMH 71
              :: :  |   : |  ::||:      |: |  || | | | || :|| || :::|| |
Db         15 DSCIGADPHKSLEELTKAIEEKSDCAEYYCQRAYAQILLQNYNDAVLDAKRSLELQPNNA 74

Qy         72 KAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVE 131
              |: ||| |   |: | :|: :   |     :   |   :| |:|::           ||
Db         75 SAFLRKGEAEFHLQNYSSAEESFRKGQMLDTSTPTFPTWIKRCEEKL----------NVS 124

Qy        132 ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDF 191
              |         :|  || :          : ||||||:| : :::::|:  | |  ::| :
Db        125 A--------EEQLTNNQQ--------ASPKYRHDWYQTESQIIITVMIKNVQKNNVHVQL 168

Qy        192 GEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSG 251
              |: |:: : :|  | |      |   |:|::   :||||||||:| | | : | ||:
Db        169 SERELTIDMNLPSGENYSLNLHLLHTILPDQSVLKVLSTKVEIKLKKTEAIRWETLEGKA 228

Qy        252 RPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIY 311
              :    |  ||      ||  ||   |:|||  |:|:||| |||:||||||: |: ||
Db        229 DSQV---KHFTPESMHKYPS--SSHYTKNWDKLVVEIKEEEKNEKLEGDAALNQLFQQIY 283
```

```
Qy        312 KDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 359
              | :::::||| |||:|| |||||||| ||| |||| :||| || ||:
Db        284 SDGNDEVKRAMNKSFMESGGTVLSTNWTDVGKKKVEVNPPDDMEWKKF 331


RESULT 14
Q8LKF8
ID   Q8LKF8      PRELIMINARY;      PRT;   211 AA.
AC   Q8LKF8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   SGT1-like protein (Fragment).
GN   Name=SGT1.2;
OS   Nicotiana benthamiana.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4100;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22155850; PubMed=12119413; DOI=10.1073/pnas.152330599;
RA   Peart J.R., Lu R., Sadanandom A., Malcuit I., Moffett P., Brice D.C.,
RA   Schauser L., Jaggard D.A.W., Xiao S., Coleman M.J., Dow M.,
RA   Jones J.D.G., Shirasu K., Baulcombe D.C.;
RT   "Ubiquitin ligase-associated protein SGT1 is required for host and
RT   nonhost disease resistance in plants.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:10865-10869(2002).
DR   EMBL; AF516181; AAM94380.1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
FT   NON_TER       1      1
FT   NON_TER     211    211
SQ   SEQUENCE   211 AA;  22945 MW;  5FC530F305121D15 CRC64;

  Query Match             32.4%;  Score 602.5;  DB 2;  Length 211;
  Best Local Similarity   57.0%;  Pred. No. 1.7e-31;
  Matches  122;  Conservative   29;  Mismatches   44;  Indels   19;  Gaps    2;

Qy         22 ATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMHKAYYRKGAAC 81
              | :||:|||    |  |:|:||||||:|||  :|||| |||||||||| | ||| ||| ||
Db          1 AVDLYTQAIAMTPKNAELFADRAQANIKLNYFTEAVVDANKAIELDPSMSKAYLRKGLAC 60

Qy         82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVEAT-------- 133
              |||||||||||||| |:| |  :||| :|:||||||||||: : | ::|: |
Db         61 IKLEEYQTAKAALETGASLAPAESRFTKLIKECDERIAEEAGELPNQSVDKTSGNVVAPP 120

Qy        134 -------VAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGVPADS 186
                    ||   :| :   |:          ::|||||::|    |||:|||||:|| :
Db        121 ASESLNNVAVDTKDAQPTVNLSYQGSA----ARPKYRHEFYQKPEEVVVTIFAKGIPAKN 176

Qy        187 VVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIP 220
              ||:|||||:||||||:|||:| | |||||| || |
Db        177 VVVDFGEQILSVSIDVPGDETYSFQPRLFGKITP 210


RESULT 15
Q8LKF9
ID   Q8LKF9      PRELIMINARY;      PRT;   211 AA.
AC   Q8LKF9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   SGT1-like protein (Fragment).
GN   Name=SGT1.1;
OS   Nicotiana benthamiana.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4100;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22155850; PubMed=12119413; DOI=10.1073/pnas.152330599;
RA   Peart J.R., Lu R., Sadanandom A., Malcuit I., Moffett P., Brice D.C.,
RA   Schauser L., Jaggard D.A.W., Xiao S., Coleman M.J., Dow M.,
RA   Jones J.D.G., Shirasu K., Baulcombe D.C.;
RT   "Ubiquitin ligase-associated protein SGT1 is required for host and
RT   nonhost disease resistance in plants.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:10865-10869(2002).
DR   EMBL; AF516180; AAM94379.1; -.
DR   InterPro; IPR007052; CS.
DR   InterPro; IPR008978; HSP20_chap.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF04969; CS; 1.
DR   Pfam; PF00515; TPR_1; 1.
DR   PROSITE; PS50005; TPR; 1.
DR   PROSITE; PS50293; TPR_REGION; 1.
FT   NON_TER       1      1
FT   NON_TER     211    211
SQ   SEQUENCE   211 AA;  22872 MW;  09D2048D8C9BE5CD CRC64;

  Query Match             32.4%;  Score 602.5;  DB 2;  Length 211;
  Best Local Similarity   57.1%;  Pred. No. 1.7e-31;
  Matches  120;  Conservative   31;  Mismatches   48;  Indels   11;  Gaps    2;

Qy         22 ATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMHKAYYRKGAAC 81
              | :||:|||    |  |:|:||||||:|||  :|||| |||||||||| | ||| ||| ||
Db          1 AVDLYTQAIAMTPKNAELFADRAQANIKLNYFTEAVVDANKAIELDPSMSKAYLRKGLAC 60

Qy         82 IKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIED- 140
              :||||||||||||| |:| |  :||| :|:||||||||||: : | ::|: |    :
Db         61 MKLEEYQTAKAALETGASLAPAESRFTKLIKECDERIAEEAGELPNQSVDKTSGNVVAPP 120

Qy        141 -KEDFTNMENTPPVIEP---------PSKPKYRHDYYNSATEVVLTIFAKGVPADSVVID 190
               |   |:   |   :|          ::|||||::|    |||:|||||:|| :|::|
Db        121 ASESLGNVAVAPKDAQPTVNLSYQGSAARPKYRHEFYQKPEEVVVTIFAKGIPAKNVIVD 180

Qy        191 FGEQMLSVSIEVPGEEPYHFQPRLFSKIIP 220
              ||||:|||||:|||:| | |||||| || |
Db        181 FGEQILSVSIDVPGDETYSFQPRLFGKITP 210


Search completed: April 20, 2005, 18:34:22
Job time : 182 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 18:28:41 ; Search time 134 Seconds
(without alignments)
895.417 Million cell updates/sec

Title: US-10-609-078-8
Perfect score: 1861
Sequence: 1 MAASDLESKAKEAFVDDDFE.....GAKKVEGSPPDGMELKKWEY 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1861	100.0	388	15	US-10-425-114-61162	Sequence 61162, A
2	1861	100.0	390	15	US-10-425-114-61204	Sequence 61204, A
3	1771.5	95.2	395	16	US-10-767-701-45515	Sequence 45515, A
4	1740	93.5	393	15	US-10-425-114-61212	Sequence 61212, A
5	1740	93.5	416	15	US-10-425-114-65836	Sequence 65836, A
6	1578	84.8	376	16	US-10-437-963-139712	Sequence 139712,
7	1263	67.9	256	15	US-10-425-114-61200	Sequence 61200, A
8	1176.5	63.2	357	15	US-10-424-599-224769	Sequence 224769,
9	1146.5	61.6	359	15	US-10-424-599-208541	Sequence 208541,
10	925.5	49.7	357	15	US-10-424-599-241114	Sequence 241114,
11	729	39.2	160	15	US-10-424-599-261373	Sequence 261373,
12	615	33.0	344	15	US-10-264-049-3093	Sequence 3093, Ap
13	603.5	32.4	223	16	US-10-437-963-146392	Sequence 146392,
14	599	32.2	168	15	US-10-424-599-224775	Sequence 224775,
15	594	31.9	365	14	US-10-287-218-12	Sequence 12, Appl
16	594	31.9	365	16	US-10-474-291-12	Sequence 12, Appl
17	471	25.3	229	16	US-10-437-963-146382	Sequence 146382,
18	464	24.9	196	10	US-09-983-802-393	Sequence 393, App
19	464	24.9	196	10	US-09-984-490-393	Sequence 393, App
20	464	24.9	196	11	US-09-973-278-503	Sequence 503, App
21	464	24.9	228	10	US-09-983-802-401	Sequence 401, App
22	464	24.9	228	10	US-09-984-490-401	Sequence 401, App
23	464	24.9	228	11	US-09-973-278-511	Sequence 511, App
24	445	23.9	413	14	US-10-032-585-7066	Sequence 7066, Ap
25	442.5	23.8	123	15	US-10-424-599-241112	Sequence 241112,
26	405	21.8	400	14	US-10-128-714-3417	Sequence 3417, Ap
27	403	21.7	123	15	US-10-424-599-218698	Sequence 218698,
28	383	20.6	478	14	US-10-128-714-8417	Sequence 8417, Ap
29	239	12.8	82	15	US-10-424-599-192478	Sequence 192478,
30	235.5	12.7	483	16	US-10-437-963-195520	Sequence 195520,
31	226.5	12.2	482	15	US-10-424-599-191442	Sequence 191442,
32	224	12.0	482	14	US-10-298-638-27	Sequence 27, Appl
33	224	12.0	485	14	US-10-298-638-10	Sequence 10, Appl
34	223	12.0	533	14	US-10-298-638-28	Sequence 28, Appl
35	219.5	11.8	163	16	US-10-767-701-41397	Sequence 41397, A
36	219.5	11.8	488	15	US-10-425-114-61206	Sequence 61206, A
37	213	11.4	101	15	US-10-424-599-183984	Sequence 183984,
38	203.5	10.9	955	9	US-09-991-496-127	Sequence 127, App
39	197.5	10.6	1427	9	US-09-874-923-97	Sequence 97, Appl
40	197.5	10.6	1427	9	US-09-991-496-97	Sequence 97, Appl
41	197.5	10.6	1641	9	US-09-874-923-96	Sequence 96, Appl
42	197.5	10.6	1641	9	US-09-991-496-96	Sequence 96, Appl
43	192	10.3	494	14	US-10-043-487-296	Sequence 296, App
44	190	10.2	590	15	US-10-425-114-55903	Sequence 55903, A
45	186	10.0	191	16	US-10-767-701-35628	Sequence 35628, A

ALIGNMENTS

RESULT 1
US-10-425-114-61162
; Sequence 61162, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61162
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3059-061-D10_FLI.pep
US-10-425-114-61162

Query Match 100.0%; Score 1861; DB 15; Length 388;
Best Local Similarity 100.0%; Pred. No. 3.1e-137;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 87

Qy         61 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 147
```

```
Qy        121 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 207

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 267

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 327

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 387

Qy        361 Y 361
              |
Db        388 Y 388
```

RESULT 2
US-10-425-114-61204
; Sequence 61204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61204
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3732-024-C2_FLI.pep
US-10-425-114-61204

Query Match 100.0%; Score 1861; DB 15; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.1e-137;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 89

Qy         61 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 149

Qy        121 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 209

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 269

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 329

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 389

Qy        361 Y 361
              |
Db        390 Y 390
```

RESULT 3
US-10-767-701-45515
; Sequence 45515, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45515
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(395)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2718_1.pep
US-10-767-701-45515

Query Match 95.2%; Score 1771.5; DB 16; Length 395;
Best Local Similarity 94.2%; Pred. No. 3.2e-130;
Matches 343; Conservative 7; Mismatches 11; Indels 3; Gaps 1;

```
Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              |||||||||||||||||||||| ||||:||||||||||:||||||||||||||||||||||
Db         32 MAASDLESKAKEAFVDDDFELAAELYTQAIDAGPATAELYADRAQAHIKLGNYTEAVADA 91

Qy         61 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 120
              |||||||| ||||||||||||||||||||||||||||||||||||||||| ||||||:|||||
Db         92 NKAIELDPTMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFTRLLKECEERIAE 151

Qy        121 ESSQAPVKNVE---ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTI 177
              |||||| ||||   |  |||:|||||  ||||||||:|||||||  ||||||||||||||
Db        152 ESSQAPAKNVEPPVAAAAATVEDKEDVANMENTPPVVEPPSKPKXSHDYYNSATEVVLTI 211

Qy        178 FAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLA 237
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        212 FAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFAKIIPEKCKYQVLSTKVEIRLA 271

Qy        238 KAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKL 297
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        272 KAEQVTWTTLDYSGRPKAVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKL 331

Qy        298 DGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELK 357
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        332 DGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVEGSPPDGMELK 391

Qy        358 KWEY 361
              ||||
Db        392 KWEY 395
```

RESULT 4
US-10-425-114-61212
; Sequence 61212, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

```
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61212
;   LENGTH: 393
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3960-013-B4_FLI.pep
US-10-425-114-61212

  Query Match             93.5%;  Score 1740;  DB 15;  Length 393;
  Best Local Similarity   93.1%;  Pred. No. 9.3e-128;
  Matches  335;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              |||||||||||||||||||| |||:||||||||||||||||||||||||||||||||||
Db         33 MAASDLESKAKEAFVDDDFELAAELYTQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 92

Qy         61 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 120
              |||| ||| |||||||||||||||||||||||||||||||| ||||| |||||||| |||
Db         93 NKAIGLDPTMHKAYYRKGAACIKLEEYQTAKAALELGSSYAPGDSRFTRLLKECDECIAE 152

Qy        121 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 180
              |||||| ||||| |||||:|||||  ||:|||||:|||||||||||||||||||||:||
Db        153 ESSQAPAKNVEAPVAATVEDKEDVANMDNTPPVVEPPSKPKYRHDYYNSATEVVLTIYAK 212

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GVPADSVVIDFGDQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 272

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
              |||||||||||||| :|||||||||||||||||:|||||||||||||||||||:||
Db        273 QVTWTTLDYSGRPKAIPQKISTPAETAPRPSYPSSKSKKDWDKLEAEVKKEEKEEKLEGD 332

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              |||||||||||||||||||| ||| |||||||||||||:| || ||||||||||||
Db        333 AALNKFFRDIYKDADEDMRRAMDKSFRESNGTVLSTNWKDVGSKTVEASPPDGMELKKWE 392


RESULT 5
US-10-425-114-65836
; Sequence 65836, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65836
;   LENGTH: 416
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3591-005-E5_FLI.pep
US-10-425-114-65836

  Query Match             93.5%;  Score 1740;  DB 15;  Length 416;
  Best Local Similarity   93.1%;  Pred. No. 1e-127;
  Matches  335;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 MAASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 60
              |||||||||||||||||||| |||:||||||||||||||||||||||||||||||||||
Db         56 MAASDLESKAKEAFVDDDFELAAELYTQAIDAGPATADLYADRAQAHIKLGNYTEAVADA 115

Qy         61 NKAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAE 120
              |||| ||| |||||||||||||||||||||||||||||||| ||||| |||||||| |||
Db        116 NKAIGLDPTMHKAYYRKGAACIKLEEYQTAKAALELGSSYAPGDSRFTRLLKECDECIAE 175

Qy        121 ESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAK 180
              |||||| ||||| |||||:|||||  ||:|||||:|||||||||||||||||||||:||
Db        176 ESSQAPAKNVEAPVAATVEDKEDVANMDNTPPVVEPPSKPKYRHDYYNSATEVVLTIYAK 235

Qy        181 GVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 240
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GVPADSVVIDFGDQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAE 295

Qy        241 QVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGD 300
              |||||||||||||| :|||||||||||||||||:|||||||||||||||||||:||
Db        296 QVTWTTLDYSGRPKAIPQKISTPAETAPRPSYPSSKSKKDWDKLEAEVKKEEKEEKLEGD 355

Qy        301 AALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWE 360
              |||||||||||||||||||| ||| |||||||||||||:| || ||||||||||||
Db        356 AALNKFFRDIYKDADEDMRRAMDKSFRESNGTVLSTNWKDVGSKTVEASPPDGMELKKWE 415


RESULT 6
US-10-437-963-139712
; Sequence 139712, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139712
;   LENGTH: 376
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_40977C.1.pep
US-10-437-963-139712

  Query Match             84.8%;  Score 1578;  DB 16;  Length 376;
  Best Local Similarity   82.1%;  Pred. No. 4.2e-115;
  Matches  307;  Conservative   27;  Mismatches   24;  Indels   16;  Gaps    3;

Qy          2 AASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADAN 61
              |||||||||| |||||||||| |||:|||:| ||||:|||||||||||||||||||||||
Db          5 AASDLESKAKAAFVDDDFELAAELYTQAIEASPATAELYADRAQAHIKLGNYTEAVADAN 64

Qy         62 KAIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEE 121
              ||||||| ||||| |||||||:|||||||||||||| |:||||||| ||:||||||||||
Db         65 KAIELDPSMHKAYLRKGAACIRLEEYQTAKAALELGYSFASGDSRFTRLMKECDERIAEE 124

Qy        122 SSQAPVKNVE-----ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLT 176
              |: |||  |      :||: :|:|:|  ||:||||::|   |||||||:||||||||||
Db        125 LSEVPVKKAEDGAAAPSVASFVEEKDDAANMDNTPPMVE--VKPKYRHDFYNSATEVVLT 182

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              |||||||::||:|||||||||||||||||||||||||||||||||| :|||||||||||
```

```
Db        183 IFAKGVPAENVVVDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKSRYQVLSTKVEIRL 242

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTP---------AETAPRPSYPSSKAKKDWDKLEAE 287
              |||||:|||:|||  :|| |||||  |         ||:| ||||||||:|||||||||
Db        243 AKAEQITWTSLDYDKKPKAVPQKIIPPVLLIALLLSAESAQRPSYPSSKSKKDWDKLEAE 302

Qy        288 VKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVE 347
              ||||||||||:|||||||||||| ||||||||||||||||||||||||||:||||
Db        303 VKKEEKEEKLEGDAALNKFFRDIYSDADEDMRRAMMKSFVESNGTVLSTNWKDVGSKKVE 362

Qy        348 GSPPDGMELKKWEY 361
              ||||||||||||||
Db        363 GSPPDGMELKKWEY 376


RESULT 7
US-10-425-114-61200
; Sequence 61200, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61200
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3689-210-C7_FLI.pep
US-10-425-114-61200

  Query Match             67.9%;  Score 1263;  DB 15;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-90;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        120 EESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFA 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 EESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFA 74

Qy        180 KGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKA 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 KGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKA 134

Qy        240 EQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 EQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDG 194

Qy        300 DAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 DAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 254

Qy        360 EY 361
              ||
Db        255 EY 256


RESULT 8
US-10-424-599-224769
; Sequence 224769, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224769
;   LENGTH: 357
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(357)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_44998C.1.pep
US-10-424-599-224769

  Query Match             63.2%;  Score 1176.5;  DB 15;  Length 357;
  Best Local Similarity   64.5%;  Pred. No. 1e-83;
  Matches  234;  Conservative   48;  Mismatches   68;  Indels   13;  Gaps    6;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              ||||| |||||| ||:::|| :| :|||   |  |||||||| :||: | ||||:||||
Db          2 ASDLELKAKEAFEDDNYDLAYDLLTQAIGLSPNNADLYADRAQVNIKVNNLTEAVSDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:|   ||| ||| |||||||||||||||:|:| | |||:|  |:|:||| |||||
Db         62 AIELNPSHSKAYLRKGTACIKLEEYQTAKAALEMGASLAPGDSKFTDLIKDCDELIAEES 121

Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPS----KPKYRHDYYNSATEVVLTIF 178
                |::   |  |  :  |     ||  |  |||:    ||||||::|    |:|:|||
Db        122 GVIPIQEESTTQGAATKAVE----AENDLP--EPPTVTVVKPKYRHEFYQKPDEMVITIF 175

Qy        179 AKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAK 238
              |||:| ||: :|||||:|||:| :| :: | |||||| |||| ||:|:|||||:|| |||
Db        176 AKGIPRDSITVDFGEQILSVTINIPCKDAYVFQPRLFGKIIPSKCRYEVLSTKIEICLAK 235

Qy        239 AEQVTWTTLDYSGRPKTVPQKIST-PAETAPRPSYPSSKAKK-DWDKLEAEVKKEEKEEK 296
              |: : ||:|::: :  || |: |  |     :|:||||| || |||||||:||||||||
Db        236 ADHIQWTSLEFN-KGSTVAQRFSVLPVARGEKPTYPSSKPKKXDWDKLEAQVKKEEKEEK 294

Qy        297 LDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMEL 356
              |||||||||||||||:||||| |||| |||||||||||||||:||: ||: ||||||||
Db        295 LDGDAALNKFFRDIYQDADEDTRRAMSKSFVESNGTVLSTNWKEVGSMKVQESPPDGMEL 354

Qy        357 KKW 359
              |||
Db        355 KKW 357


RESULT 9
US-10-424-599-208541
; Sequence 208541, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 208541
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_30341C.1.pep
```

```
US-10-424-599-208541

  Query Match             61.6%;  Score 1146.5;  DB 15;  Length 359;
  Best Local Similarity   62.8%;  Pred. No. 2.3e-81;
  Matches  230;  Conservative   55;  Mismatches   66;  Indels   15;  Gaps    5;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              ||||:||||||||:| ||||  :| ||||     |  |:  ||||||:|||  |:|||||||||
Db          2 ASDLEAKAKEAFVEDHFELAVDLLSQAIHLEPNKAEFYADRAQANIKLNNFTEAVADANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              ||||:|  :  |||  |||  ||:|||||:||||||:|:|  :   :||||  |:||||:  |||||
Db         62 AIELNPSLPKAYLRKGTACMKLEEYETAKAALEVGASLSPDNSRFATLIKECDKLIAEES 121

Qy        123 SQAPVKNVEATVAATIEDKEDFT--NMENTPPVIEPP----SKPKYRHDYYNSATEVVLT 176
                 |:   :||          || |   :::     ::| |     :||||||::|      :||:|
Db        122 YTPPIIEEKATT-------EDVTPKDVQQQDDLLEKPTVTVTKPKYRHEFYQKHDQVVVT 174

Qy        177 IFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRL 236
              |||| :| :|: :||||:||||| | ||:  |  ||||||  ||||   |:|:|||||:||||
Db        175 IFAKKIPKESITVDFGEQILSVSINVTGEDTYVFQPRLFGKIIPSNCQYEVLSTKIEIRL 234

Qy        237 AKAEQVTWTTLDYSGRPKTVPQKISTPAET-APRPSYPSSKAKKDWDKLEAEVKKEEKEE 295
              ||||  :   ||:|:::  |       |  |::|     :   :|:||||   ::      :   ||:||||  |:|
Db        235 AKAEPIHWTSLEFT-RDSVVQQRVSALSVTGGEKPTYPSSKRCENTTEGEAQVKKEGKDE 293

Qy        296 KLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGME 355
              |||||||||||||:||:|||||  ||||  |||||||||||||||||:||  |||||||||||||
Db        294 KLDGDAALNKFFREIYQDADEDTRRAMKKSFVESNGTVLSTNWKEVGTKKVEGSPPDGME 353

Qy        356 LKKWEY 361
              ||||||
Db        354 LKKWEY 359
```

```
RESULT 10
US-10-424-599-241114
; Sequence 241114, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 241114
;   LENGTH: 357
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(357)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_59754C.1.pep
US-10-424-599-241114

  Query Match             49.7%;  Score 925.5;  DB 15;  Length 357;
  Best Local Similarity   53.3%;  Pred. No. 4.5e-64;
  Matches  194;  Conservative   53;  Mismatches  104;  Indels   13;  Gaps    5;

Qy          3 ASDLESKAKEAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANK 62
              |: || |||||| ||:| ||  :|||:||     |     |:|:||||||||||   :||||:|||
Db          2 ATALEKKAKEAFFDDEFGLAVDLYSEAIRLDPNDANLFADRAQAHIKLNAFTEAVSDANK 61

Qy         63 AIELDPMMHKAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEES 122
              :|:|:| :  ||| ||   |||||:|| ||| ||: |:::|  ||||| |:::||  |||
Db         62 SIQLNPSLPKAYLRKATACIKLQEYHTAKVALQNGAAFAQDDSRFANLIQQCDRCIAEXE 121
```

```
Qy        123 SQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGV 182
              |          |   :|  :       :  :  :      :  |      ||      :|  :  |      |      :  :
Db        122 SS------ERLLAPCLPHLSNKKSWDEXKKLKETVCFPKRMEPHYTNPCEGSNTTXHRNL 175

Qy        183 PADSVVIDFGEQMLSVS---IEVPGEEPYHFQPRLFS-KIIPEKCKYQVLSTKVEIRLAK 238
              |           ::|  :        ::|   |    |   |  ||     |||   |: :|||||:||||||
Db        176 DQDGAPFMMFHKLLQLMHQLLQVKDEHSQHVFPCLFGYMIIPNNCRVEVLSTKIEIRLAK 235

Qy        239 AEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAK-KDWDKLEAEVKKEEKEEKL 297
              || :  ||:|:|         |:|   |:  |   :  |  ||||  |  :  ||||||||:||||||||||
Db        236 AEAINWTSLEYG--KNTLPPIINRPIVQSERASYPSPKPRTKDWDKLEAQVKKEEKEEKL 293

Qy        298 DGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELK 357
              ||||||| |||||::|||||||||| |||:|||||||||||:||:||:|||||||||:||||
Db        294 DGDAALNKLFRDIYQNADEDMRRAMSKSFLESNGTVLSTDWKEVGSKKVEGSPPEGMELK 353

Qy        358 KWEY 361
              ||||
Db        354 KWEY 357
```

```
RESULT 11
US-10-424-599-261373
; Sequence 261373, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 261373
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_78041C.1.pep
US-10-424-599-261373

  Query Match             39.2%;  Score 729;  DB 15;  Length 160;
  Best Local Similarity   89.2%;  Pred. No. 3.6e-49;
  Matches  140;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         97 GSSYASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEP 156
              ||||| ||||| |||||||| |||||||| ||||| ||||:||||| ||:|||||:||
Db          4 GSSYAPGDSRFTRLLKECDECIAEESSQAPAKNVEAPVAATVEDKEDVANMDNTPPVVEP 63

Qy        157 PSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFS 216
              |||||| ||||||||||||||:|||||||||||| |:||||||||||||||||| ||| |
Db         64 PSKPKYTHDYYNSATEVVLTIYAKGVPADSVVIDCGDQMLSVSIEVPGEEPYHFXPRLLS 123

Qy        217 KIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRP 253
              ||||||||| ||||||||||||||||||||||||||||
Db        124 KIIPEKCKYHVLSTKVEIRLAKAEQVTWTTLDYSGRP 160
```

```
RESULT 12
US-10-264-049-3093
; Sequence 3093, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3093
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-049-3093

  Query Match             33.0%;  Score 615;  DB 15;  Length 344;
  Best Local Similarity   38.7%;  Pred. No. 8.6e-40;
  Matches  135;  Conservative   62;  Mismatches  116;  Indels   36;  Gaps    5;

Qy         12 EAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMH 71
              :| :|:| : | |  ::|::  |  |  |  ||  || ||||  ||||| |::||:|
Db         31 DALIDEDPQAALEELTKALEQKPDDAQYYCQRAYCHILLGNYCVAVADAKKSLELNPNNS 90

Qy         72 KAYYRKGAACIKLEEYQTAKAALELGSSYASGDSRFARLLKECDERIAEESSQAPVKNVE 131
             |  |||     : |  |       |    | |: |:  :| | |  |:  |:: |
Db         91 TAMLRKGICEYHEKNYAAALETFTEGQKLDSADANFSVWIKRCQE--AQNGSESEV---- 144

Qy        132 ATVAATIEDKEDFTNMENTPPVIEPPSKPKYRHDYYNSATEVVLTIFAKGVPADSVVIDF 191
                          :|:            : | ::|:| : ::||:|:  | |  : | ::|
Db        145 ------------WTH------------QSKIKYDWYQTESQVVITLMIKNVQKNDVNVEF 180

Qy        192 GEQMLSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSG 251
              |: ||  :::|  | |: :  |   ||||:  ::|||||:||:| | | | |  |:  |
Db        181 SEKELSALVKLPSGEDYNLKLELLHPIIPEQSTFKVLSTKIEIKLKKPEAVRWEKLEGQG 240

Qy        252 RPKTVPQKISTPAETAPRPSYPSSKA-KKDWDKLEAEVKKEEKEEKLDGDAALNKFFRDI 310
               |  | ::      :  ||||    ::||||  |:|:||| |||:||||||: |: |
Db        241 DVPTPKQFVAD-----VKNLYPSSSPYTRNWDKLVGEIKEEEKNEKLEGDAALNRLFQQI 295

Qy        311 YKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKW 359
              | |  ::::||| |||:|| |||||||| ||| :||| :||| || ||:
Db        296 YSDGSDEVKRAMNKSFMESGGTVLSTNWSDVGKRKVEINPPDDMEWKKY 344


RESULT 13
US-10-437-963-146392
; Sequence 146392, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146392
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_47020C.1.pep
US-10-437-963-146392

  Query Match             32.4%;  Score 603.5;  DB 16;  Length 223;
  Best Local Similarity   47.5%;  Pred. No. 3.8e-39;
  Matches  125;  Conservative   45;  Mismatches   44;  Indels   49;  Gaps    7;

Qy        101 ASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSKP 160
              | ||:|||:| || |::   ::: ||    |   ||:                    ||
Db          8 APGDARFAKLAKELDDKAPPKAAAAPPAASEDQAAAS------------------SGKP 48

Qy        161 KYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKIIP 220
              ||:||:|| |:|||:|:||| | |: | :||||||||||:||                 :|
Db         49 KYKHDFYNGASEVVVTVFAKSVAAEHVSVDFGEQMLSVSVEV----------------VP 92

Qy        221 EKCKYQVLSTKVEIRLAKA-EQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKAKK 279
              ||:| :| :|:|:||||| ||||||:|:|: :      |::  | |       .:: |
Db         93 AKCRYSILPSKIEVRLAKADEQVTWTSLEYTSKAN---NKLAATATT-------TTRKKV 142

Qy        280 DWDKLEAEVKKEEKEEKLD-GDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNW 338
              |||||||||||:||::|     :|:||: :| : |||||||:|||:|||   ||||:|
Db        143 DWDKLEAEVKKEEEEEEVDTATPVVNRFFQQMYGNGDEDMRRAIMKSYVES--YVLSTDW 200

Qy        339 KDVGAKKVEGSPPDGMELKKWEY 361
              ||||:||:| | |:|||| ||||
Db        201 KDVGSKKIEASAPEGMELHKWEY 223


RESULT 14
US-10-424-599-224775
; Sequence 224775, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224775
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(168)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_45001C.1.pep
US-10-424-599-224775

  Query Match             32.2%;  Score 599;  DB 15;  Length 168;
  Best Local Similarity   65.9%;  Pred. No. 5.7e-39;
  Matches  116;  Conservative   20;  Mismatches   22;  Indels   18;  Gaps    2;

Qy        196 LSVSIEVPGEEPYHFQPRLFSKIIPEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKT 255
              |||:| :| :: | |||||| |||| ||:|:|||||:|| ||| : : ||:||::
Db          1 LSVTINIPCKDAYVFQPRLFGKIIPSKCRYEVLSTKIEICLAKTDHIQWTSLDFN----- 55

Qy        256 VPQKIST----------PAETAPRPSYPSSKAKKDWDKLEAEVKKEEKEEKLDGDAALNK 305
                | ||               |||||||  :||||:||:||||||:||||||||||
Db         56 ---KASTDILHRGLMLLQLLEVXRPSYPSSKQTRDWDKIEAQVKKEEKDEKLDGDAALNK 112

Qy        306 FFRDIYKDADEDMRRAMMKSFVESNGTVLSTNWKDVGAKKVEGSPPDGMELKKWEY 361
              |||:||:||||| |||| ||||||||||||||:||:|||||| ||||||||||
Db        113 FFREIYQDADEDTRRAMKKSFVESNGTVLSTNWKEVGSKKVEGSAPDGMELKKWEY 168


RESULT 15
US-10-287-218-12
; Sequence 12, Application US/10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
```

```
;  APPLICANT:  AZIMZAI, Yalda; AU-YOUNG, Janice K.
;  APPLICANT:  BATRA, Sajeev; BAUGHN, Mariah R.
;  APPLICANT:  BECHA, Shanya D.; BOROWSKY, Mark L.
;  APPLICANT:  BUFORD, Neil; DING, Li
;  APPLICANT:  ELLIOTT, Vicki S.; EMERLING, Brooke M.
;  APPLICANT:  GANDHI, Ameena R.; GIETZEN, Kimberly J.
;  APPLICANT:  GRIFFIN, Jennifer A.; HAFALIA, April J.A.
;  APPLICANT:  HONCHELL, Cynthia D.; LAL, Preeti G.
;  APPLICANT:  LEE, Soo Yeun; LU, Dyung Aina M.
;  APPLICANT:  ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
;  APPLICANT:  REDDY, Roopa; SANJANWALA, Madhu, M.
;  APPLICANT:  TANG, Y. Tom; WALIA, Narinder K.
;  APPLICANT:  WANG, Yu-mei, E.; WARREN, Bridget A.
;  APPLICANT:  XU, Yuming; YANG, Junming
;  APPLICANT:  YAO, Monique G.; YUE, Henry
;  APPLICANT:  ZEBARJADIAN, Yeganeh
;  TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
;  FILE REFERENCE: PI-0417 USA
;  CURRENT APPLICATION NUMBER: US/10/287,218
;  CURRENT FILING DATE:  2002-10-31
;  PRIOR APPLICATION NUMBER: PCT/US02/11152
;  PRIOR FILING DATE: 2002-04-05
;  PRIOR APPLICATION NUMBER: US 60/349,705
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: US 60/295,263
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/295,340
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/293,727
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/291,846
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/291,662
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: US 60/287,228
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/286,820
;  PRIOR FILING DATE: 2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/283,294
;  PRIOR FILING DATE: 2001-04-11
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 12
;   LENGTH: 365
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030198975A1 1543848CD1
US-10-287-218-12

  Query Match             31.9%;  Score 594;  DB 14;  Length 365;
  Best Local Similarity   36.2%;  Pred. No. 4.1e-38;
  Matches  138;  Conservative   68;  Mismatches  107;  Indels   68;  Gaps    9;

Qy          12 EAFVDDDFELATELYSQAIDAGPATADLYADRAQAHIKLGNYTEAVADANKAIELDPMMH 71
               :| :|:| : | |  ::|::  |  |  |  ||  || ||||  ||||| |::||:|
Db          20 DALIDEDPQAALEELTKALEQKPDDAQYYCQRAYCHILLGNYCVAVADAKKSLELNPNNS 79

Qy          72 KAYYRKG---------AACI-------KLE----EYQTAKAALEL-----------GSS 99
                 |  |||          || :        ||:      ::  :| |:|           :
Db          80 TAMLRKGICEYHEKNYAAALETFTEGQKLDIETGFHRVGQAGLQLLTSSDPPALDSQSAG 139

Qy         100 YASGDSRFARLLKECDERIAEESSQAPVKNVEATVAATIEDKEDFTNMENTPPVIEPPSK 159
                   |: |:  :| | |  |:  |:: |                 :|:            :
Db         140 ITGADANFSVWIKRCQE--AQNGSESEV----------------WTH-----------Q 169

Qy         160 PKYRHDYYNSATEVVLTIFAKGVPADSVVIDFGEQMLSVSIEVPGEEPYHFQPRLFSKII 219
                 | ::|:| : ::||:|:  | |  : | ::| |: ||  :::|  | |: :  |   ||
Db         170 SKIKYDWYQTESQVVITLMIKNVQKNDVNVEFSEKELSALVKLPSGEDYNLKLELLHPII 229

Qy         220 PEKCKYQVLSTKVEIRLAKAEQVTWTTLDYSGRPKTVPQKISTPAETAPRPSYPSSKA-K 278
                 ||:  ::|||||:||:| | | | |:  |   |  | ::       :  ||||
Db         230 PEQSTFKVLSTKIEIKLKKPEAVRWEKLEGQGDVPTPKQFVAD-----VKNLYPSSSPYT 284

Qy         279 KDWDKLEAEVKKEEKEEKLDGDAALNKFFRDIYKDADEDMRRAMMKSFVESNGTVLSTNW 338
                 ::|||  |:|:||| |||:|||||: |: || |  ::::||| |||:|| ||||||||
Db         285 RNWDKLVGEIKEEEKNEKLEGDAALNRLFQQIYSDGSDEVKRAMNKSFMESGGTVLSTNW 344

Qy         339 KDVGAKKVEGSPPDGMELKKW 359
                 ||| :||| :||| || ||:
Db         345 SDVGKRKVEINPPDDMEWKKY 365


Search completed: April 20, 2005, 18:38:15
Job time : 136 secs
```

FILE 'HOME' ENTERED AT 16:05:29 ON 03 MAY 2005

=> file biosis caplus caba agricola
=> s rar1
L1 165 RAR1

=> duplicate remove l1
L2 80 DUPLICATE REMOVE L1 (85 DUPLICATES REMOVED)

=> d ti 1-25

L2 ANSWER 1 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Calcium-independent phospholipase A2 is regulated by a novel protein kinase C in human coronary artery endothelial cells

L2 ANSWER 2 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
TI Mammalian CHORD-containing protein 1 is a novel heat shock protein 90-interacting protein.

L2 ANSWER 3 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI **RAR1**, ROR1, and the actin cytoskeleton contribute to basal resistance to Magnaporthe grisea in barley

L2 ANSWER 4 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Barley Rom1 reveals a potential link between race-specific and nonhost resistance responses to powdery mildew fungi

L2 ANSWER 5 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI The atypical resistance gene, RPW8, recruits components of basal defence for powdery mildew resistance in Arabidopsis

L2 ANSWER 6 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Pesticidal cyanine dye derivatives

L2 ANSWER 7 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
TI **RAR1** positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance.

L2 ANSWER 8 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on STN
TI Arabidopsis RIN4 negatively regulates disease resistance mediated by RPS2 and RPM1 downstream or independent of the NDR1 signal modulator and is not required for the virulence functions of bacterial type III effectors AvrRpt2 or AvrRpm1.

L2 ANSWER 9 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Cutting edge transcriptome analysis: It's all about design

L2 ANSWER 10 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on
TI Molecular chaperone Hsp90 associates with resistance protein N and its signaling proteins SGT1 and **Rar1** to modulate an innate immune response in plants.

L2 ANSWER 11 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Prothrombin overexpressed in postnatal neurons requires blood factors for activation in the mouse brain

L2 ANSWER 12 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on
TI Signaling requirements and role of salicylic acid in HRT- and rrt-mediated resistance to turnip crinkle virus in Arabidopsis.

L2 ANSWER 13 OF 80 CAPLUS COPYRIGHT 2005 ACS on STN
TI Mammalian CHORD-containing protein 1 is a novel heat shock protein 90-interacting protein

L2 ANSWER 14 OF 80 BIOSIS COPYRIGHT (c) 2005 The Thomson Corporation on